

# Track Reconstruction Efficiency and Analysis of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ at the LHCb Experiment

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde  
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Michel De Cian

von

Brügg BE

Promotionskomitee

Prof. Dr. Ulrich Straumann (Vorsitz)

Dr. Olaf Steinkamp

Dr. Nicola Serra

Zürich, 2013

## Abstract

The LHCb experiment at the Large Hadron Collider is a particle physics experiment dedicated to the investigation of so-called  $B$  mesons. To track long-living charged particles traversing the detector, LHCb comprises three tracking stations and a dipole magnet. The efficiency to reconstruct a long-living particle is of crucial importance for many physics analyses. A novel method to access this track reconstruction efficiency is presented and discussed in detail. Rare decays of  $B$ -mesons are a prospective way to search for physics beyond the Standard Model of particle physics, as new physics can enter at the same level as Standard Model physics. The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is an ideal laboratory for such searches, as its four-particle final state gives rise to many angular distributions which can be measured. One quantity of particular interest is the zero-crossing point of the forward-backward asymmetry  $A_{\text{FB}}$ , as it can be predicted theoretically with a small uncertainty. The first measurement of this zero-crossing point and the evaluation of its experimental uncertainty is presented, using the collision data collected during 2011 with LHCb. The measurement is in agreement with the prediction of the Standard Model. The analyses of precisely predicted quantities in the  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  decay is then expanded to four observables called  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  which are measured for the first time in an experiment. While  $P'_4$ ,  $P'_6$  and  $P'_8$  show a good agreement with the prediction by the Standard Model,  $P'_5$  shows a deviation from the Standard Model prediction by about four standard deviations.

## Zusammenfassung

Das LHCb-Experiment am Large Hadron Collider ist ein Teilchenphysik-Experiment für die Erforschung der sogenannten B-Mesonen. Langlebige, geladene Teilchen werden mit Hilfe von drei Spurrekonstruktions-Stationen und einem Dipolmagneten rekonstruiert. Die Effizienz dieser Spurrekonstruktion ist von entscheidender Bedeutung für viele Analysen von Teilchenzerfällen. Eine neuartige Methode um diese Effizienz zu berechnen wird vorgestellt und im Detail diskutiert. Seltene Zerfälle von B-Mesonen sind besonders geeignet für die Suche nach Physik jenseits des Standardmodells, da «neue» Physik gleichberechtigt mit Physik des Standardmodells auftauchen kann. Der Zerfall  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  ist ein ideales «Labor» für solche Analysen, da sein Vier-Teilchen Endzustand die Messung vieler Winkelverteilungen erlaubt. Eine Grösse von besonderem Interesse ist der Nulldurchgang der Vorwärts-Rückwärts-Asymmetrie  $A_{\text{FB}}$ , da die theoretische Vorhersage nur kleine Unsicherheiten aufweist. Die erste Messung dieses Nulldurchgangs und die Bestimmung der experimentellen Unsicherheit wird vorgestellt. Die Analyse beruht auf den Kollisions-Daten von LHCb aus dem Jahre 2011. Das Resultat stimmt mit der Vorhersage des Standardmodells überein. Die Analyse wird dann auf vier weitere Observablen im Zerfall  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  ausgebaut, die  $P'_4$ ,  $P'_5$ ,  $P'_6$  und  $P'_8$  genannt werden. Auch diese Grössen haben eine präzise theoretische Vorhersage und wurden zum ersten Mal in einem Experiment gemessen. Die Observablen  $P'_4$ ,  $P'_6$  und  $P'_8$  zeigen eine gute Übereinstimmung mit der Vorhersage der Standardmodell.  $P'_5$  jedoch weicht von der Vorhersage um etwa vier Standardabweichungen ab.

*«What we don't understand, we explain to each other»,*  
J. Robert Oppenheimer



# Contents

<b>Preface</b>	<b>11</b>
<b>1. The Standard Model of particle physics: A brief overview</b>	<b>13</b>
1.1. The right stuff . . . . .	13
1.2. The constituents of the Standard Model . . . . .	13
1.2.1. Quarks . . . . .	14
1.2.2. Leptons . . . . .	15
1.2.3. Bosons . . . . .	16
1.3. The formalism of the Standard Model . . . . .	16
1.3.1. Flavour changing neutral currents . . . . .	19
1.3.2. Operator product expansion . . . . .	20
1.4. The Standard Model and beyond . . . . .	21
1.5. $b$ quark production at hadron colliders . . . . .	22
<b>2. The experiment</b>	<b>27</b>
2.1. CERN . . . . .	27
2.2. The Large Hadron Collider . . . . .	28
2.2.1. Layout and Design . . . . .	28
2.2.2. LHC performance . . . . .	31
2.3. The LHCb experiment . . . . .	32
2.3.1. The magnet . . . . .	36
2.3.2. The Vertex Locator . . . . .	36
2.3.3. The Tracker Turicensis . . . . .	38
2.3.4. The Inner Tracker . . . . .	41
2.3.5. The Outer Tracker . . . . .	43
2.3.6. The Ring Imaging Cherenkov Detectors . . . . .	43
2.3.7. The calorimetry system . . . . .	45
2.3.8. The muon system . . . . .	48
2.3.9. The trigger system . . . . .	50
<b>3. Track reconstruction in LHCb</b>	<b>55</b>
3.1. The anatomy of a flight path . . . . .	55

## Contents

3.2.	Passage of particles through matter . . . . .	55
3.2.1.	Electromagnetic interactions . . . . .	56
3.2.2.	Hadronic interactions . . . . .	57
3.2.3.	Multiple Scattering . . . . .	58
3.3.	Representation of the flight path of a particle . . . . .	58
3.3.1.	Parametrisation of a track . . . . .	58
3.3.2.	Kalman filtering for the track fit . . . . .	59
3.4.	Tracks types in LHCb . . . . .	60
3.5.	Pattern recognition algorithms . . . . .	62
3.5.1.	Forward Tracking . . . . .	62
3.5.2.	Track Matching . . . . .	63
3.5.3.	Additional pattern recognition algorithms . . . . .	64
3.5.4.	Final track selection . . . . .	65
3.6.	Adding TT hits to long tracks . . . . .	65
3.6.1.	The TT hit adding algorithm . . . . .	65
3.6.2.	Tuning of parameters . . . . .	69
3.6.3.	Performance . . . . .	71
3.6.4.	TT hit maps . . . . .	73
<b>4.</b>	<b>Tracking efficiency in LHCb</b>	<b>77</b>
4.1.	Approaches to measure the tracking efficiency . . . . .	77
4.2.	Construction of MuonTT tracks . . . . .	78
4.2.1.	Muon standalone tracks . . . . .	78
4.2.2.	Formation of muon standalone tracks . . . . .	79
4.2.3.	Momentum estimate . . . . .	80
4.2.4.	MuonTT tracks . . . . .	82
4.3.	Reconstruction of dimuon resonances in collider data . . . . .	84
4.3.1.	Constructing a dimuon resonance . . . . .	85
4.4.	Prerequisites for a tracking efficiency measurement . . . . .	86
4.4.1.	Association . . . . .	86
4.4.2.	General crosschecks . . . . .	87
4.5.	Tracking efficiency for high transverse momentum tracks . . . . .	90
4.5.1.	Event selection . . . . .	92
4.5.2.	Checks on simulated data . . . . .	92
4.5.3.	Checks on collision data . . . . .	94
4.5.4.	Results . . . . .	95
4.6.	Tracking efficiency for low transverse momentum tracks . . . . .	100
4.6.1.	Event selection . . . . .	100
4.6.2.	Data samples . . . . .	101
4.6.3.	Fit model . . . . .	102
4.6.4.	Checks on simulated data . . . . .	102

4.6.5.	Checks on collision data . . . . .	104
4.6.6.	Reweighting . . . . .	104
4.6.7.	Impact of track quality cut . . . . .	105
4.6.8.	Results for the Long Method . . . . .	105
4.6.9.	Calculation of uncertainties . . . . .	105
4.7.	Alternative method to measure the tracking efficiency for long tracks . . . . .	112
4.7.1.	VELO tracking efficiency . . . . .	112
4.7.2.	T station tracking efficiency . . . . .	113
4.8.	Combined results for low transverse momentum tracks . . . . .	114
4.8.1.	Additional systematic effects . . . . .	115
4.8.2.	Results . . . . .	115
4.8.3.	Bremsstrahlung and hadronic interactions . . . . .	115
<b>5.</b>	<b>Theoretical introduction to <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math></b>	<b>117</b>
5.1.	$B^0 \rightarrow K^{*0} \mu^+ \mu^-$ : An intriguing decay . . . . .	117
5.2.	Decay topology . . . . .	117
5.2.1.	Definition of the angles . . . . .	119
5.2.2.	Differential decay rate . . . . .	121
5.2.3.	Definition of related observables . . . . .	123
5.2.4.	Reducing the number of observables . . . . .	124
5.2.5.	Extracting the $S_i$ observables using a counting experiment . . . . .	127
5.2.6.	Introducing an S-wave system . . . . .	128
5.3.	Theoretical uncertainties and clean observables . . . . .	131
5.3.1.	Definition of the $P'_i$ observables . . . . .	132
5.4.	Theoretical predictions . . . . .	132
5.5.	Measurements of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ by CDF and the $B$ -factories . . . . .	133
<b>6.</b>	<b>Selecting <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math></b>	<b>137</b>
6.1.	Selection . . . . .	137
6.1.1.	Trigger selection . . . . .	137
6.1.2.	Central offline selection («stripping selection») . . . . .	137
6.1.3.	Preselection . . . . .	138
6.1.4.	Multivariate selection . . . . .	139
6.1.5.	Charmonium Vetoes . . . . .	140
6.1.6.	Peaking backgrounds . . . . .	141
6.2.	Corrections to the simulation . . . . .	144
6.2.1.	Impact parameter resolution . . . . .	144
6.2.2.	Muon identification . . . . .	144
6.2.3.	Systematic uncertainty for the muon identification . . . . .	148
6.2.4.	Kaon identification . . . . .	150
6.2.5.	Further corrections and checks to the simulation . . . . .	152

## Contents

6.3.	Correction for detector acceptance . . . . .	152
<b>7.</b>	<b>Results of the «first» angular analysis of <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math></b>	<b>155</b>
<b>8.</b>	<b>Determination of the zero-crossing point of <math>A_{\text{FB}}</math></b>	<b>157</b>
8.1.	Extraction of the zero-crossing point with a linear fit . . . . .	157
8.2.	The «unbinned counting» technique . . . . .	158
8.2.1.	Choice of the order of the polynomial . . . . .	160
8.2.2.	Independence of fit range . . . . .	161
8.3.	Choice and determination of probability distribution functions for the fit .	161
8.4.	Checks on simulated data . . . . .	163
8.4.1.	Simulated data after reconstruction, signal only . . . . .	163
8.4.2.	Simulated data after reconstruction, signal and background . . . .	164
8.4.3.	Determination of the zero crossing point . . . . .	165
8.5.	Determination of the zero-crossing point in collision data . . . . .	166
8.5.1.	The numerical result . . . . .	166
8.5.2.	Bootstrapping . . . . .	168
8.5.3.	Fit-and-Toy technique . . . . .	168
8.6.	Systematic uncertainties . . . . .	171
8.7.	Goodness-of-Fit tests . . . . .	172
8.8.	Final result . . . . .	172
<b>9.</b>	<b>Measuring the <math>S_i</math> and <math>P'_i</math> observables in <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math></b>	<b>175</b>
9.1.	Cross-checks with toy studies . . . . .	175
9.1.1.	Pulls . . . . .	175
9.1.2.	Values for $F_L$ and $A_T^{(2)}$ . . . . .	178
9.1.3.	Cross-check with counting experiments . . . . .	178
9.2.	Cross-checks with simulated data . . . . .	180
9.2.1.	Comparison between a fit and a counting experiment for the $S_i$ observables . . . . .	181
9.2.2.	Extracting the $S_i$ observables with and without folding . . . . .	182
9.3.	Fit on collision data . . . . .	184
9.3.1.	Calculation of the statistical uncertainty . . . . .	184
9.3.2.	Fit results . . . . .	185
9.3.3.	Comparison with counting experiment . . . . .	187
9.4.	Estimating the S-wave contribution . . . . .	187
9.5.	Systematic uncertainties . . . . .	191
9.6.	Experimental results and comparison with theoretical predictions . . . .	191
9.7.	Further cross-checks . . . . .	192
9.7.1.	Comparison between $S_i$ and $P'_i$ observables . . . . .	192
9.7.2.	Further checks for the third bin of $P'_5$ . . . . .	194

9.8. Conclusion . . . . .	195
<b>10. The End</b>	<b>197</b>
<b>A. Linear track fit</b>	<b>199</b>
<b>B. Calculation of uncertainty with sideband subtraction</b>	<b>201</b>
<b>C. Extraction of a reference value of the zero-crossing point of <math>A_{\text{FB}}</math> in simulation</b>	<b>203</b>
<b>D. Derivation of formulae for the <math>S_i</math> and <math>P'_i</math> observables</b>	<b>205</b>
D.1. Transformations for $P'_4/S_4$ . . . . .	206
D.2. Transformations for $P'_5/S_5$ . . . . .	207
D.3. Transformations for $P'_6/S_7$ . . . . .	207
D.4. Transformations for $P'_8/S_8$ . . . . .	208
<b>E. S-wave contribution to <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math></b>	<b>211</b>
<b>F. Pulls for <math>A_T^{(2)}</math> and <math>F_L</math> for <math>P'_4</math></b>	<b>213</b>
<b>G. Systematic uncertainties for <math>P'_4, P'_5, P'_6, P'_8</math></b>	<b>217</b>
<b>Acknowledgements</b>	<b>233</b>
<b>Curriculum Vitae</b>	<b>235</b>



# Preface

In Alfred Hitchcock's film «Rear Window» the photographer L. B. Jeffries, played by James Stewart, is bound to a wheelchair and stuck in his Greenwich Village apartment after an accident. Out of boredom he starts observing his neighbours through the rear window of his flat. At some point, he discovers the strange behaviour of a salesman, one of his neighbours, and comes to the conclusion that he must have murdered his wife.

Analysing the smallest constituents of matter and their interactions sometimes feels like being in Jeffries role. The human body is used to detecting scales of metres not femtometres – the eye can detect photons in the energy range of eV, not GeV. One is faced with a huge gap in the scale of perception, much like Jeffries only sees the events from the other side of the apartment block. To observe the details, to finally conclude, we need to resort to sophisticated technical apparatuses; the telescope lens in «Rear Window» becomes the particle accelerator, the camera the particle detector. In contrast to Hitchcock's thriller, research in particle physics does not follow a scripted plot: the story does not end after two hours. There is no Grace Kelly bridging the gap between the events and the observer that delivers the crucial piece of evidence. One is left with a nature that does answer the questions one asks only indirectly. And no final truth is within eyeshot. It seems like every layer uncovered reveals another mystery, another inconsistency.

The situation then faintly resembles the one of Sisyphus, the tragic hero of greek mythology, who is burdened to roll a stone up a hill, only to let it roll down again and start his task all over. In a famous essay, the French writer Albert Camus took Sisyphus as the symbol for the absurdity of life which does not give answers to existential questions. This applies to particle physics as well: it is absurd in the sense that the researcher is let alone with his results. He deals with a nature that does not follow any discernible path. And there is no final understanding, no deepest level of knowledge which could be attained. And no hope for a final meaning.

The question then remains why this basic research is undertaken at all. Camus concludes that one must imagine Sisyphus happy. He is aware of the futility of his undertaking as he knows that he will not overcome his fate. But he takes his life in his own hands when he struggles with the stone on the slope. The motivation for research is not to reach a final understanding, to find the meaning of it all. It's about charting the unknown possibilities. It's the pleasure of finding out things.





# 1. The Standard Model of particle physics: A brief overview

## 1.1. The right stuff

The Standard Model of Particle Physics (SM) describes the interplay between elementary particles and the electromagnetic, the weak and the strong forces.<sup>1</sup> It is essentially a merger between two theories: the «Glashow-Weinberg-Salam» (GWS) model describing the electromagnetic and the weak interactions; and «Quantum Chromodynamics» (QCD), describing the strong interaction. The Standard Model describes the experimental findings to a high precision in many different regimes, ranging from low energies like  $\beta$ -decay up to very high energies at hadron colliders. However, the Standard Model is not a «theory of everything». It is not complete in the sense that it does not describe all phenomena observed in nature (*e.g.* the composition of dark matter), and that it has theoretical shortcomings (*e.g.* the quadratic divergence of the Higgs mass, see for example Ref. [2]). And maybe most importantly, it does not incorporate gravity.

This chapter gives a short overview of the constituents and the formulation of the Standard Model, quickly highlights its shortcomings and describes the production of  $b$  quarks, which are of fundamental interest for this thesis. A theoretical overview of the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  whose analysis will be discussed in this thesis is given in Chap. 5.

The introduction to the Standard Model in the following pages is by no means complete. For a more thorough overview consider Ref. [3]. An introduction to QCD is given in Refs. [4] [5] and a treatment of the GWS model (aka «Electroweak Theory») is given in Ref. [6]. Different aspects of the Standard Model and beyond are discussed in Ref. [7].

## 1.2. The constituents of the Standard Model

In nature, particles are either realised as «fermions», particles with half-integer spin, or «bosons», particles with integer spin. All the quarks and leptons, the particles normally associated to what is called «matter» are fermions. The bosons are the force-carriers, the particles that are exchanged between the fermions and amongst each other. For the electromagnetic force they «transmit» the attracting or repelling force which is responsible for

---

<sup>1</sup>The term «particle» in physics is a «physical system that has no continuous degrees of freedom except for its momentum» [1].

bound states. One particular boson, the Higgs boson, is responsible for the particles to acquire mass.

### 1.2.1. Quarks

Quarks are massive fermions. The lightest of them are thought to be the elementary constituents<sup>2</sup> of what makes up the nuclei of atoms, *i.e.* protons and neutrons. They can be classified as «up»- or «down»-type according to their fractional electric charge, which is either  $-1/3$  (down-type) or  $+2/3$  (up-type).<sup>3</sup> The Standard Model incorporates six different quarks, named «up», «down», «charm», «strange», «top», «bottom/beauty», and the corresponding antiquarks. The masses of the quarks are a free parameter of the Standard Model and ranges from several  $\text{MeV}/c^2$  (up and down quark) up to about  $172 \text{ GeV}/c^2$  for the top quark. Quarks can be classified in three families. The up-type and the corresponding down-type quarks always build a doublet of weak isospin  $T$ , the quantity coupling the quarks to the weak interaction<sup>4</sup>. The order of the families is given by the mass of the quarks by convention. In addition to their electrical and «weak» charge, quarks also carry colour charge with three possible colours denoted as  $R$ ,  $G$  or  $B$ . Quarks are therefore affected by all three interactions in the Standard Model. Their properties are summarised in Table 1.1 and 1.2, their quantum numbers are shown in Table 1.6.

For this thesis the bottom (b) quark is of special interest. With a mass of about  $4.2 \text{ GeV}/c^2$  in the  $\overline{\text{MS}}$  mass scheme,<sup>5</sup> it is the second heaviest quark in the Standard Model and has a very rich decay structure leading to many different final states. Since its discovery in 1977 in the  $\Upsilon(1S)$  resonance [9] and the discovery of the  $B$  meson [10] [11], analyses of particles containing a  $b$  quark have allowed the extraction of many observables of the Standard Model, *e.g.* oscillations of neutral  $B$  mesons [12] or determination of CKM-matrix elements [13] (see later). The  $b$  quark is the heaviest quark known to form bound states at the present time.

Table 1.1.: Classification of quarks according to their electrical charge and family.

Charge	1. family	2. family	3. family
$+2/3$	up (u)	charm (c)	top (t)
$-1/3$	down (d)	strange (s)	bottom (b)

<sup>2</sup>The quarks are considered «pointlike» *i.e.* no substructure has been found yet.

<sup>3</sup>The charge is always given in fractions of the fundamental charge,  $1.602 \cdot 10^{-19} \text{C}$ .

<sup>4</sup>More precisely: only the left-handed quarks have a non vanishing weak isospin

<sup>5</sup>The masses of the quarks cannot be measured directly (except the mass of the top) and the calculation is dependent on the renormalisation scheme.  $\overline{\text{MS}}$  is one particular scheme and it is used for the masses quoted in Ref. [8].

Table 1.2.: Masses of the quarks. The numbers are taken from Ref. [8]. Only the top quark mass comes from a direct measurement. Note the comment about the calculation of quark masses in Ref. [8].

Name	Mass
down	$4.1 - 5.8 \text{ MeV}/c^2$
up	$1.7 - 3.3 \text{ MeV}/c^2$
strange	$101^{+29}_{-21} \text{ MeV}/c^2$
charm	$1.27^{+0.07}_{-0.09} \text{ GeV}/c^2$
bottom	$4.19^{+0.18}_{-0.06} \text{ GeV}/c^2$
top	$172 \pm 0.9 \pm 1.3 \text{ GeV}/c^2$

### 1.2.2. Leptons

Leptons are fermions and, like quarks, can be classified in three families<sup>6</sup>. They can be divided in charged leptons and neutral leptons, called neutrinos. However, while the charged leptons have mass, the neutrinos are massless<sup>7</sup>. They can only interact weakly, while the charged leptons can also interact electromagnetically. The mass of the charged leptons ranges from  $511 \text{ MeV}/c^2$  for the electron up to  $1777 \text{ MeV}/c^2$  for the  $\tau$ -lepton. In the Standard Model, the lepton family number is conserved *i.e.* the number of leptons belonging to a certain family has to stay constant throughout a process. This is in contrast to the quark families, where this is not the case. The leptons and some of their properties are listed in Table 1.3 and 1.4 and their quantum numbers are given in Table 1.6.

Table 1.3.: Classification of leptons according to their electrical charge and family.

Charge	1. family	2. family	3. family
-1	electron (e)	muon ( $\mu$ )	tau ( $\tau$ )
0	electron-neutrino ( $\nu_e$ )	muon-neutrino ( $\nu_\mu$ )	tau-neutrino ( $\nu_\tau$ )

<sup>6</sup>For the GWS mechanism to be mathematically consistent, the number of lepton and quark families have to agree, see Ref. [14].

<sup>7</sup>Experimental results show that neutrinos do actually have mass, albeit as small one, see for example Refs. [15] and [16]. However, their mass is normally not explained with the same mechanism as the mass for the other fermions (the Higgs mechanism), but with models that are beyond the Standard Model. If not explicitly stated, neutrinos are assumed massless in the following.

Table 1.4.: Masses of the leptons. The numbers for the charged leptons are taken from Ref. [8] and the upper limits for the neutrinos from Ref. [17].

Name	Mass
electron	$0.511 \text{ MeV}/c^2$
muon	$105.7 \text{ MeV}/c^2$
tau	$1777 \text{ MeV}/c^2$
electron-neutrino	$< 2 \text{ eV}/c^2$
muon-neutrino	$< 0.19 \text{ MeV}/c^2$
tau-neutrino	$< 18.2 \text{ MeV}/c^2$

### 1.2.3. Bosons

The Standard Model incorporates five fundamental bosons. Four act as force-carriers and have spin 1: the photon ( $\gamma$ ) for the electromagnetic force, the  $W$  and  $Z^0$  bosons for the weak force and the gluon ( $g$ ) for the strong force. Photons are electrically neutral and cannot interact with each other. The  $Z^0$  bosons are also electrically neutral, while  $W$  bosons are charged — both carriers of the weak force can interact with each other via triple or quartic boson couplings. The gluons carry colour charge and undergo self-interactions. The presence of self-interactions of gluons and their absence for photons manifests itself in the different behaviour of the electromagnetic and the strong force as a function of the distance. While the strong coupling becomes stronger for larger distances, the electromagnetic force becomes weaker. Gluons exist in eight types, each having a different combination of colour and anti-colour.<sup>8</sup> The last boson is the Higgs boson. It has spin 0 and is responsible for the mass of the  $W$  and  $Z^0$  bosons as well as the masses of the quarks and the charged leptons. The five bosons and their properties are summarised in Table 1.5.

## 1.3. The formalism of the Standard Model

The Standard Model is formulated using the framework of quantum field theory where the quarks, leptons and bosons are represented via quantum fields. These fields obey certain relations, namely they are invariant under local transformations of a gauge group,  $SU(3) \times SU(2) \times U(1)$  in the case of the Standard Model. This behaviour is called «local gauge invariance» and results in five distinct quantum fields for the fermions. In Table 1.6 the fermion fields with some quantum numbers are listed. An important property of these fields is their «chirality»: they are either left- or right-handed. The chirality can be obtained

<sup>8</sup>The ninth gluon is colourless and therefore does not couple to anything.

Table 1.5.: Fundamental bosons in the Standard Model and some of their properties.

Name	Electrical charge	Spin	Mass [ GeV/ $c^2$ ]	Interaction strength
Photon $\gamma$	0	1	0	$\alpha \sim 1/137$
$W$ boson	$\pm 1$	1	80.4	$G_F \sim 10^{-5} \frac{1}{\text{GeV}}$
$Z^0$ boson	0	1	90.2	
Gluon (g)	0	1	0	$\alpha_S \sim 0.1^9$
Higgs (H) boson	0	0	126 <sup>10</sup>	-

by using projectors:

$$\psi_L = P_L \psi = \frac{1}{2}(1 - \gamma_5)\psi \quad (1.1)$$

$$\psi_R = P_R \psi = \frac{1}{2}(1 + \gamma_5)\psi, \quad (1.2)$$

where  $\psi$  is a fermion field,  $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$  and  $\gamma_i$  is a Dirac  $\gamma$ -matrix. While the electromagnetic and the strong interactions do not distinguish between left- and right-handed fermions, the weak interaction with the  $W$  and  $Z^0$  bosons only couples to left-handed particles. The right handed particles consequently have a zero weak-isospin  $T$ , see Table 1.6.

The interaction of the fermions with the bosonic fields and the dynamics of the bosonic fields itself are given by the following Lagrangian:

$$\mathcal{L}_{\text{gauge}} = \sum_{j=1,\dots,3} \sum_{\psi=Q_L^j, \dots, E_R^j} \bar{\psi} i \not{D} \psi \quad (1.3)$$

$$- \frac{1}{4} \sum_{a=1,\dots,8} G_{\mu\nu}^a G_{\mu\nu}^a - \frac{1}{4} \sum_{a=1,\dots,3} W_{\mu\nu}^a W_{\mu\nu}^a - \frac{1}{4} B_{\mu\nu} B_{\mu\nu}, \quad (1.4)$$

where  $j$  runs over the three quark and lepton flavours,  $\psi$  runs over the quark and lepton singlets and doublets as described in Table 1.6,  $\not{D}$  is the covariant derivative,  $G_{\mu\nu}^a$  is the

<sup>9</sup>The coupling strength depends on the energy which is mostly due to the self-coupling of the gluons. For high energies the coupling is weak and the quarks basically behave like free particles («asymptotic-freedom»). For low energies ( $\mathcal{O}(1 \text{ GeV})$ ) the quarks become strongly bound. The value given here is evaluated at the  $Z^0$  mass. The electromagnetic coupling only shows little variation as a function of the energy.

<sup>10</sup>The resonance found at about 126 GeV/ $c^2$  by the ATLAS [18] and CMS [19] collaborations is consistent with being the Higgs boson. However, no definite statement about the precise nature of the discovered particle can be given at the time of this thesis.

Table 1.6.: Listing of all fermions in the Standard Model (excluding the anti-particles). The subscript  $L$  stands for left-handed while  $R$  means right-handed.  $T$  is the weak isospin,  $T_3$  the third component of the weak-isospin,  $Y$  the weak hypercharge and  $Q$  the electrical charge in units of the fundamental charge.  $T_3$ ,  $Y$  and  $Q$  are connected via  $Q = T_3 + \frac{Y}{2}$ , the Gell-Mann–Nishijima formula.

Symbol	Family			$T$	$T_3$	$Y$	$Q$
$L_L$	$\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau,L} \\ \tau \end{pmatrix}$	$1/2$	$+1/2$ $-1/2$	$-1$ $-1$	$0$ $-1$
$E_R$	$e_R$	$\mu_R$	$\tau_R$	$0$	$0$	$-2$	$-1$
$Q_L$	$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s'_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b'_L \end{pmatrix}$	$1/2$	$+1/2$ $-1/2$	$1/3$ $1/3$	$+2/3$ $-1/3$
$U_R$	$u_R$	$c_R$	$t_R$	$0$	$0$	$+4/3$	$+2/3$
$D_R$	$d_R$	$s_R$	$b_R$	$0$	$0$	$-2/3$	$-1/3$

field-strength tensor for the gluon field (with  $a$  running over all eight gluons), and  $W_{\mu\nu}^a$  and  $B_{\mu\nu}$  are the fields of the electroweak theory. The first part of the Lagrangian corresponds to the interaction between the fermionic matter and the bosonic fields and is essentially a Dirac equation for massless fermions. The second part corresponds to the field strength of the bosons.

As the fermions, leptons and the  $W$  and  $Z^0$  bosons are not massless in nature, a mechanism has to be introduced to let them acquire mass. However, the introduction of a Dirac mass term (like  $m\psi\bar{\psi}$ ) would destroy gauge-invariance, which is not desired. A different approach therefore has to be followed, which is achieved by the Higgs-mechanism [20] [21] [22] [23] [24]. It introduces a complex scalar field  $\phi$ , the Higgs field, with a non-zero vacuum expectation value  $v$  throughout space. If the potential  $V(\phi)$  is chosen such that its minimum is not symmetric with respect to  $\phi$ , mass terms can be generated without breaking gauge invariance. This is called «spontaneous symmetry breaking», as the ground state is not in a symmetric position of the potential  $V(\phi)$ . For the quarks, this results in the introduction of «Yukawa-couplings»<sup>11</sup> which lets them acquire mass via interactions with the Higgs field:

$$\mathcal{L}_{\text{Yukawa}} = Y_d^{ij} \bar{Q}_L^i \phi D_R^j + Y_u^{ij} \bar{Q}_L^i \tilde{\phi} U_R^j (+h.c.), \quad (1.5)$$

with  $\phi$  the Higgs field and  $Y_{d,u}^{ij}$  the coupling constants for the mass. The mass of the quarks are then  $m_q = Y_q \frac{v}{\sqrt{2}}$  and therefore proportional to their coupling to the Higgs

<sup>11</sup>The situation is similar for charged leptons.

field.<sup>12</sup> The matrices  $Y_{d,u}^{ij}$  are not diagonal; to diagonalise them four independent matrices would be needed. However, only three of them can be chosen freely, hence only one of the  $Y$  matrices can be diagonalised. Normally the one with the up-type quarks is chosen, which leaves the one with the down-type quarks non-diagonal. This results then in:

$$Y_u = \mathbb{I} \cdot \begin{pmatrix} u \\ c \\ t \end{pmatrix} \quad Y_d = \mathbb{I} \cdot \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.6)$$

with  $\mathbb{I}$  the identity matrix and  $V_{CKM}$  the Cabbibo-Kobayashi-Maskawa (CKM) [25] matrix. It relates the flavour eigenstates  $(d', s', b')$  to the mass eigenstates  $(d, s, b)$ .

The CKM matrix with all its elements is:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (1.7)$$

The matrix is unitary. The elements off the diagonal show a strong hierarchical order:  $|V_{us}|$  and  $|V_{cd}|$  are about 0.22,  $|V_{cb}|$  and  $|V_{ts}|$  of order  $4 \cdot 10^{-2}$  and  $|V_{ub}|$  and  $|V_{td}|$  of order  $5 \cdot 10^{-3}$ . As the matrix is unitary and one global phase is not observable, four free parameters remain. Three are the quark mixing angles and one is a complex phase. This complex phase gives rise to « $\mathcal{CP}$  violation» in the Standard Model, *i.e.* the different behaviour of particles and anti-particles in the weak interaction.

Another consequence of the CKM-matrix and the magnitude of its elements is that hadrons with a  $b$  quark component have a long lifetime. As the decay to a top quark is forbidden by the mass difference, the  $b$  quark can only decay to lighter flavours, which is suppressed by the CKM-element. This is commonly referred to as «Cabibbo suppression». The effect is even more pronounced in processes which, for example, involve the transition of a  $b$  to a  $u$  quark.

### 1.3.1. Flavour changing neutral currents

At first glance, transitions of the sort  $b \rightarrow s$  or  $s \rightarrow d$  seem possible by exchanging a  $Z^0$  boson. However, in the Standard Model this is not the case on tree-level<sup>13</sup>, see Fig. 1.1. The «GIM-Mechanism» [26] states that flavour changing neutral currents (FCNCs) can only occur on levels higher than tree-level if the quarks are arranged in doublets as in the Standard Model. The reason is that the neutral current is diagonal in flavour space when having six quarks in three doublets. The GIM mechanism could also explain why the branching fraction of the decay  $K^0 \rightarrow \mu^+ \mu^-$  was much smaller than expected. The

<sup>12</sup>That also means that this mechanism has no predictive power for the quark masses, as the Yukawa couplings are not known *a priori*.

<sup>13</sup>Tree-level stands for first order in perturbation theory.

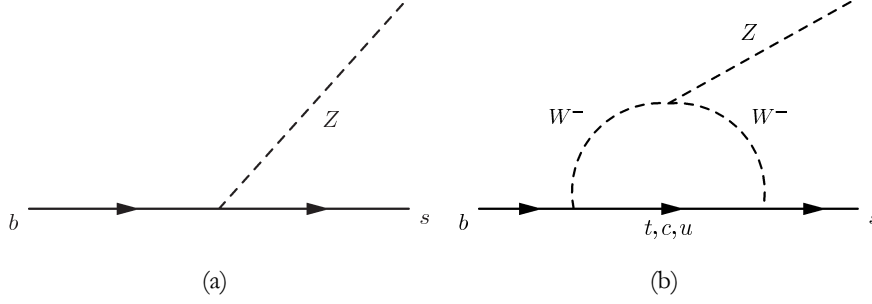


Figure 1.1.: Feynman diagrams for the flavour changing neutral current  $b \rightarrow s$ . Diagram (a) is not allowed in the SM, FCNCs in the lowest order are only realised as loop diagrams as depicted in diagram (b).

authors proposed the existence of a new quark, the  $c$  quark, whose existence would lead to destructive interference and therefore to a vanishing of this decay mode. In reality, the finite mass difference between the  $s$  and the  $c$  quark breaks the GIM mechanism; hence the branching fraction of  $K^0 \rightarrow \mu^+ \mu^-$  is only strongly suppressed but not zero.

The only possibility then to study FCNCs is to examine loop-processes. However, these processes are more rare than tree-level ones, as they involve at least two charged flavour-changing currents. One of these currents is between two quark families which is suppressed by the corresponding element of the CKM-matrix. An example for such a decay is the process  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  which will be discussed in detail in the Chaps. 5, 6, 7, 8 and 9.

### 1.3.2. Operator product expansion

Processes of B meson decays involve two different energy scales: the electroweak scale, which is defined by the  $W$  mass, and  $\Lambda_{\text{QCD}}$ , governing the energy scale of the hadronisation. Treating these two scales together leads to complications in the calculation of the decay amplitudes. A solution is to integrate out the heavy fields of the Standard Model (the  $W$  and  $Z^0$  bosons and the top quark) leaving an «effective» theory where only the light fields appear. This is similar to the Fermi-theory of the  $\beta$ -decay, where only a four-point interaction was considered and the  $W$  was integrated out as the existence of the  $W$  and  $Z^0$  bosons were not known at that time. [27].

The matrix element for a quark decaying from some initial state  $i$  to some final state  $f$  in the effective theory is:

$$\mathcal{M}_{i \rightarrow f} = \langle f | \mathcal{H}_{\text{eff}} | i \rangle = -4 \frac{G_F}{\sqrt{2}} \lambda_{CKM} \sum_i \mathcal{C}_i(\mu) \langle f | \mathcal{O}_i(\mu) | i \rangle. \quad (1.8)$$

$\mathcal{H}_{\text{eff}}$  is the effective Hamiltonian,  $G_F$  is the Fermi constant,  $\lambda_{CKM}$  the CKM matrix ele-



ments corresponding to the decay and  $\mu$  is the renormalisation scale. The scale  $\mu$  separates the problem in two parts: one of short distance effects, incorporated into the  $\mathcal{C}_i(\mu)$  and one of long distance effects, incorporated in the  $\mathcal{O}_i(\mu)$ . The  $\mathcal{C}_i(\mu)$  are called Wilson coefficients. If new physics with masses above the masses of the Standard Model would exist, it could alter the Wilson coefficients and would therefore be visible in an experiment as a deviation from the calculation.

In principle there are an infinite number of operators  $\mathcal{O}_i(\mu)$ . However not every operator contributes to every decay. Furthermore terms with  $\mathcal{O}(\mu^2/M_W^2)$  and higher can be neglected as they are suppressed by the large mass of the  $W$ . For the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ , the following operators are important in the Standard Model:

$$\begin{aligned}\mathcal{O}_7 &= \frac{e}{16\pi^2} m_b (\bar{s} \sigma^{\mu\nu} P_R b) F_{\mu\nu} \\ \mathcal{O}_9 &= \frac{e^2}{16\pi^2} (\bar{s} \gamma^\mu P_L b) (\bar{\ell} \gamma_\mu \ell) \\ \mathcal{O}_{10} &= \frac{e^2}{16\pi^2} (\bar{s} \gamma^\mu P_L b) (\bar{\ell} \gamma_\mu \gamma_5 \ell)\end{aligned}\tag{1.9}$$

with  $e$  the fundamental electrical charge,  $F_{\mu\nu}$  the electromagnetic field tensor,  $\sigma^{\mu\nu} = [\gamma^\mu, \gamma^\nu]$  and  $P_{R,L} = (1 \pm \gamma_5)$ . The symbols  $b$  and  $s$  denote the fields for the bottom and the strange quark,  $\ell$  the field for a lepton. The operators  $\mathcal{O}_9$  and  $\mathcal{O}_{10}$  both have the same initial and final states, but  $\mathcal{O}_9$  treats vector currents, while  $\mathcal{O}_{10}$  treats axial-vector currents. The right-handed counterpart of all these operators also exist and are denoted with  $\mathcal{O}'_7$ ,  $\mathcal{O}'_9$  and  $\mathcal{O}'_{10}$ , they exchange  $P_R \leftrightarrow P_L$ . Additionally, also scalar or tensor operators can be formulated. The operators for Eqs. 1.9 are visualised in Fig. 1.2.

## 1.4. The Standard Model and beyond

Despite its remarkable success, the Standard Model is not believed to be the final theory explaining all observable phenomena in the field of particle physics. It is rather considered an effective theory that is valid at the electroweak scale. Some of the reasons are listed in the following:

- **The mystery of dark matter:** It was observed a long time ago by Zwicky [28] that the rotation of galaxies for large radial distances is too fast to be in equilibrium between gravitational and centrifugal force. It can however be explained by postulating «dark matter» in the halo of galaxies, for which more and more evidence has been found in the past years, see for example Ref. [29]. Dark matter is believed to be at most weakly interacting, apart from its gravitational interaction. In the Standard Model, no candidate for this type of matter exists. The mass of the neutrinos and their density in the universe is too small to explain the effect.

- **The riddle of baryon asymmetry:** While almost no anti-baryonic matter is observed in the Universe, baryonic matter exists in abundance. This asymmetry is unexplained, as equal amounts of matter and anti-matter should have been created at the «Big Bang»<sup>14</sup> Three necessary conditions need to be fulfilled [31] [32] to generate the observed particle-antiparticle asymmetry, of which  $\mathcal{CP}$  violation is one. However, although a certain amount of  $\mathcal{CP}$  violation is incorporated into the Standard Model with the CKM mechanism, this amount is far too small to account for the observed asymmetry.
- **The fine-tuning problem:** Although being more of an aesthetic problem than a physical one, the question is still important: Why does the Higgs mass remain relatively low when higher order loop-contributions are divergent and should make the mass infinite, or at least as large as the Planck-scale. In the Standard Model the Higgs mass can only be made finite by precisely fine-tuning tree-level and loop-contributions. Although this is not forbidden in the theory, the explanation is somewhat unsatisfactory.
- **Number of parameters in flavour physics:** Finally it has to be noted that many parameters in the Standard Model which are needed as an input relate to flavour physics: Six quark masses, three masses for the electron-type leptons, four parameters for the CKM matrix and one parameter to allow for non-perturbative  $\mathcal{CP}$  violation in QCD. The rest of the model only needs five input parameters.<sup>15</sup>

A possible extension of the Standard Model is Supersymmetry (SUSY) which doubles the amount of particles by adding a fermionic partner to every boson and a bosonic partner to every fermion in the Standard Model. Within SUSY the fine-tuning problem does not occur as the introduction of loop corrections with bosons would cancel the divergence. Furthermore it provides a dark matter candidate in the form of the lightest supersymmetric particle (LSP) which often is assumed to be the neutralino (a state which is a mixture between the superpartners of the photon, the  $Z^0$  boson and the Higgs boson). More information on SUSY can be found in Refs. [7] [33].

Although having attracted a lot of attention, SUSY is not the only model of new physics. Many further models for physics beyond the Standard Model exist. They are partly covered in Ref. [7].

## 1.5. $b$ quark production at hadron colliders

There are different production mechanisms for  $b$  quarks in proton-proton collisions. For leading-order (LO), the contributing processes are quark-antiquark annihilation ( $q\bar{q} \rightarrow b\bar{b}$ )

<sup>14</sup>The Big Bang theory describes the early development of the universe [30]. However, at the time of this thesis there also exists a highly successful US sitcom with the same name.

<sup>15</sup>A possible choice are:  $\alpha_s$ ,  $\alpha_{em}$ , the weak mixing angle  $\sin^2 \theta_W$ , the Higgs mass and the  $W$  or  $Z^0$  mass.

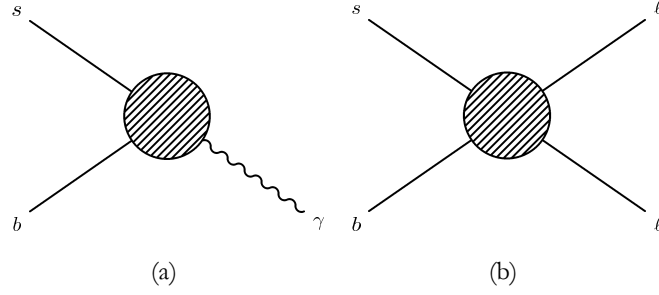


Figure 1.2.: Feynman diagrams for the operators  $\mathcal{O}_7$  (a),  $\mathcal{O}_9$  and  $\mathcal{O}_{10}$  (b).

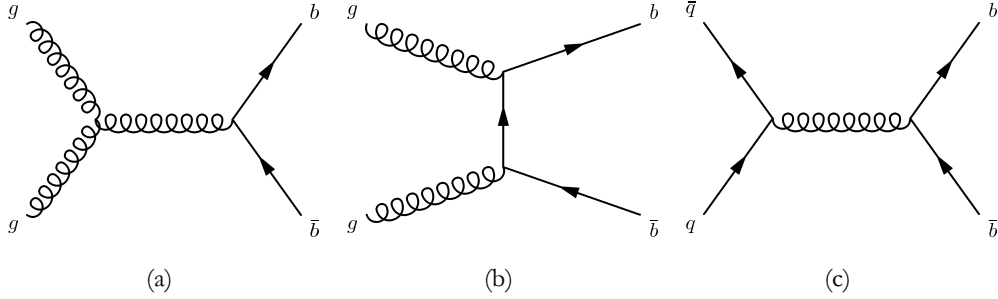


Figure 1.3.: Feynman diagrams for leading order for  $b\bar{b}$  production in proton-proton collisions. (a) and (b) show gluon-gluon fusion processes while (c) shows quark-antiquark annihilation processes.

and gluon-gluon fusion ( $gg \rightarrow b\bar{b}$ ), commonly referred to as pair creation. The corresponding Feynman diagrams are shown in Fig. 1.3. In next-to-leading order (NLO), gluon-splitting and flavour-excitation come into play, see Fig. 1.4. The contribution of the leading-order processes with respect to the total  $b$ -cross section decreases with increasing energy. For centre-of-mass energies of 7 TeV or 14 TeV at the LHC (see next chapter), the dominating process is flavour-excitation. The different contributions as implemented in the event-generator PYTHIA 6.4 [34] are shown in Fig. 1.5.

$b\bar{b}$  quark pairs are often created with a large boost and therefore tend to fly along the axis of one incoming proton. Furthermore there is a strong correlation between the  $b$  and the  $\bar{b}$  quark which causes them both to end up in the forward- or backward direction. This is visualised in Fig. 2.5 and played an important role in the design of the LHCb experiment, cf. Chap. 2. The  $b\bar{b}$  cross section has been measured in the forward region at the LHCb

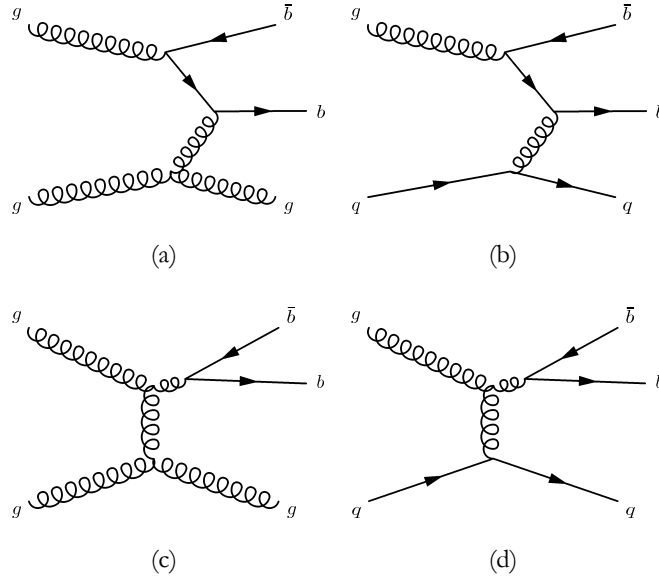


Figure 1.4.: Feynman diagrams for next-to-leading order for  $b\bar{b}$  production in proton-proton collisions. (a) and (b) show flavour excitation processes while (c) and (d) show gluon splitting processes.

experiment in the pseudorapidity<sup>16</sup> interval  $2 < \eta < 6$ <sup>17</sup> and is  $(75.3 \pm 5.4 \pm 13.0) \mu\text{b}$  [35]. Quarks do not exist as unconfined (*i.e.* free) objects and can only be observed as bound states in the form of mesons or baryons. At some point the individual quarks must fragment into colourless bound objects. These fractions are difficult to predict theoretically as the fragmentation happens in the non-perturbative regime of QCD. They can however be measured experimentally *e.g.* by measuring the relative production of different  $b$ -hadron species [36] [37]. The average results of the fragmentation fractions from different experiments are presented in Table 1.7. However it should be noted that fragmentation fractions can depend on quantities like the energy or the transverse momentum, the quoted averages may therefore not be universally applicable.

<sup>16</sup>Pseudorapidity is defined as  $\eta = -\log \tan \frac{\theta}{2}$  with  $\theta$  being the angle between the particle in question and the beam axis.

<sup>17</sup>Note that an extrapolation to  $4\pi$  is not trivial and has large uncertainties. Therefore it is not quoted here.

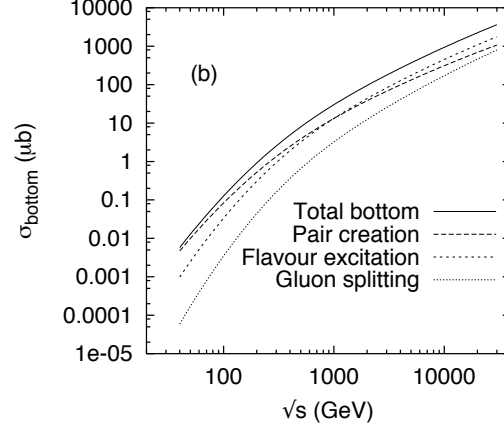


Figure 1.5.:  $b\bar{b}$  cross section at hadron colliders with different contributions due to pair creation, flavour excitation and gluon splitting. Figure taken from Ref. [38].

Table 1.7.: Fragmentation fractions for different hadrons containing a  $b$  quark. For the first row, the  $B^0$  fraction and the  $B^+$  fraction are set equal and the number only applies for one type of  $B$  meson and not for both together. The numbers are taken from Ref. [39].

Fraction	Value
$B^0$ or $B^+$ fraction	$0.401 \pm 0.007$
$B_s^0$ fraction	$0.107 \pm 0.0005$
$b$ -baryon fraction	$0.091 \pm 0.015$



## 2. The experiment

This chapter gives a «tour d’horizon» of the experimental environment and the apparatus in whose contexts the work presented in this thesis is embedded. It starts with a short description of CERN and the LHC, followed by a summary of the subdetectors of the LHCb experiment.

### 2.1. CERN

CERN, the European Organization for Nuclear Research, has its foundations after the second world war in «scientists’ search for new ways of acquiring large-scale equipment, and the statesmen’s search for new ways of common interest in which a joint effort could be made to produce tangible manifestations of European unity» [40]. In 1952, a provisional organisation called CERN (Conseil Européen pour la Recherche Nucléaire) was founded to plan the build-up of a research laboratory for nuclear research; soon after Geneva was chosen as its site. The laboratory finally was accepted by the people of the Canton of Geneva via a referendum a little later. In 1954, the organisation officially came into being with today’s full name, however, the abbreviation CERN was kept.

Research started with the first accelerator in 1957, a synchro-cyclotron, which was followed shortly afterwards by the PS (Proton Synchrotron), that accelerated the first beams on the 24<sup>th</sup> of November 1959. It is still in operation today. Since then, CERN has seen an ever increasing number of large accelerators. The «Intersecting Storage Rings» (ISR), the world’s first proton-proton accelerator, started operation in 1971, followed five years later by the «Super Proton Synchrotron» (SPS). The «Large Electron Positron Collider» (LEP) accelerated electrons and positrons from 1989 to 2000; its dismantling gave way to the last addition of the «family», the «Large Hadron Collider» (LHC) which saw the first collisions in 2009. Along with the increasing size of the accelerators, the number of experiments and buildings on the main site of CERN in Meyrin (since the construction of the ISR straddling the boarder between Switzerland and France) and in Prévessin grew more and more. Today, CERN is one of the largest research institutions worldwide, having about 10’000 visiting scientists and representing 608 universities and 113 nationalities.

## 2.2. The Large Hadron Collider

### 2.2.1. Layout and Design

The Large Hadron Collider (LHC) [41] [42]<sup>1</sup> is considered the largest scientific instrument in the world. It is a proton-proton collider with a circumference of 26.7 km with a design energy of 7 TeV per beam and housed in the former LEP tunnel.

The first studies for the LHC were already made in the 80s, during the construction of LEP. It was planned from the beginning to reuse large parts of the infrastructure, mainly the LEP tunnel. This tunnel consists of eight arcs with a length of 2.8 km and eight straight sections with a length of 500 m each and lies between 45 m and 170 m below the surface. The tunnel has an inclination of 1.42% with respect to the horizontal to allow for an easier civil engineering.<sup>2</sup> These given dimensions set the boundaries for the LHC. To achieve the design energy of 7 TeV, the bending magnets which keep the protons on track need to have a field strength of 8.3 Tesla, 60% more than previously achieved in other machines. This required the usage of superconducting magnets cooled with superfluid helium at 1.9K. It has to be pointed out that this implied the construction of a large cooling infrastructure, as about 80 tons of superfluid helium has to be maintained at this very low temperature. The next limit was the diameter of 3.8 m of the LEP tunnel, which did not allow two separate beam pipes for both beams (like in the ISR) but asked for a more compact solution. It was found by placing both beam pipes in a single cryostat, where the two pipes are only separated by 19 cm.

Although the LEP tunnel was reused for the LHC, two new transfer lines to transport the protons from the SPS to the LHC had to be constructed. The old ones (used for electrons

---

<sup>1</sup>The LHC shares its acronym with «Les Horribles Cernettes», the first band with a webpage.

<sup>2</sup>«In view of its dimensions, there was no way to avoid locating the main tunnel of LEP in two very different geological zones. Most of the tunnel, i.e. about 23 km, has been excavated in the molasses of the Lemanic basin, which is composed of a tertiary formation of consolidated fluvial and marine deposits of alpine origin. The remainder of the tunnel is situated in the piedmont region of the westernmost Secondary Period (Mesozoic) chain of the Jura.

In the Lemanic plane the molasses do not generally occur at the surface, but lie beneath moraine deposits consisting of gravels, sands and loams which can contain ground water. Such surface layers may give rise to problems for tunnel excavation since they do not lend themselves to mechanized boring methods and are not sufficiently stable for CERN accelerators requirements. [...]

The whole region has been subjected to severe folding and as a result there are a lot of cracks. The Jura chain has many fissures, either running parallel to its axis, resulting in intense erosion centered on the summit fault or combined in two transverse directions. The massif has thus been segmented into successive blocks separated by faults.

As a result of exploratory test borings, the positioning of LEP was modified to avoid crossing the Jura as much as possible. The final positioning included only 3 km in the piedmont region with a maximum rock over of 170 m. To avoid a deep depression in the moraine detected in the eastern part of the project near Geneva-Cointrin airport, the main tunnel plane was inclined by 1.42° from the horizontal.» (taken from Ref. [43])



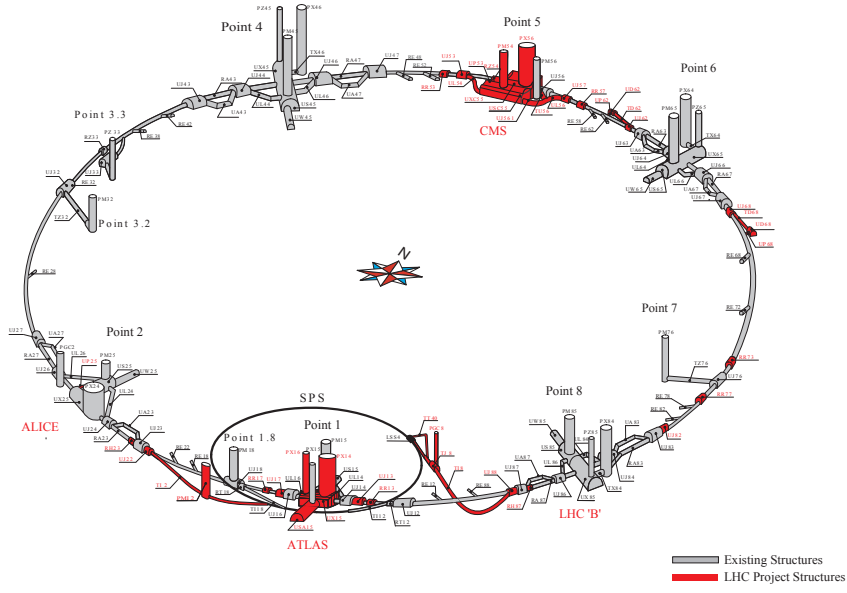


Figure 2.1.: Overview of the underground infrastructure of the LHC at the different access points. Red are the structures that were constructed for the LHC and its experiments, white are the structures that already existed from LEP (not to scale). Figure taken from Ref. [45].

and positrons) had a too small curvature for protons with 450 GeV. The full situation with all underground structures is shown in Fig. 2.1. To house the four main experiments at the LHC, four collision points had to be chosen. While ATLAS was placed at Point 1 close to the CERN main site, CMS was given the more remote location at Point 5. For both collision points, new caverns had to be built. ALICE was given Point 2 and LHCb is located at Point 8, the former home of the DELPHI experiment [44] at LEP. Not only the locations of the four experiments are different but also their scope: ATLAS and CMS are so called «General Purpose Detectors» for high transverse momentum ( $p_T$ ) physics; ALICE investigates primordial states of matter like the «quark-gluon plasma»; LHCb is a dedicated b-physics experiment, focusing on CP-violation and rare decays of beauty and charm hadrons, but also performing analyses in the field of electroweak physics and other topics.

During the circulation of the beam in the LHC for many hours, a «halo» of particles builds up around the center of the beam. These particles are off-center due to non-linearities in the magnetic field or due to interaction of both beams. A particle that hits the wall of the vacuum chamber would be dangerous for the machine and especially the magnets (leading

to a quench<sup>3</sup>), therefore blocks can be moved into the beam pipe to remove the halo. This collimation system is located at Point 3 and Point 7.

To accelerate the protons, Radio Frequency (RF) cavities are situated at Point 4. These cavities operate at a frequency of 400 MHz and accelerate the protons using a longitudinally oscillating electric field. To achieve a constant acceleration over many turns of the protons, the frequency of the RF must be precisely tuned to the revolution frequency of the LHC. Additionally the phase shift of the field must be adjusted to keep the protons on track with the rising magnetic field in the dipole magnets. Furthermore, the RF does not only accelerate the protons but also corrects for a longitudinal spread in the proton distribution, by giving protons arriving too late an additional kick while decelerating protons arriving too early<sup>4</sup>. The RF system relies on superconducting technology to avoid losses in the walls of the cavity when building the electrical field.

To allow a safe removal of the beam from the accelerator, a beam-dump system is in place at Point 6. It enables the quick firing of a kicker magnet, which steers the beam away from its normal path and deflects it into a 700 m long evacuated pipe using another magnet system. The beam is then «diluted» by sweeping it across the surface of a block of graphite, which serves as a beam dump.

It has to be noted that the LHC is only the last piece of a complicated chain of devices to produce, bunch and accelerate the protons<sup>5</sup> until they finally collide inside one of the four collision points. The full accelerator chain is shown in Fig. 2.2. The protons are extracted as the nucleus of hydrogen atoms with an energy of 50 keV and guided to a linear accelerator (LINAC), where their energy is increased to 50 MeV. They are then injected into a booster synchrotron which increases their energy to 1.4 GeV until they finally arrive at the Proton Synchrotron (PS). In this machine the protons are not only accelerated but also grouped into train of bunches, a structure that is kept until the beams finally collide in the LHC. After leaving the PS, these bunch-trains are injected into the SPS, where they are again accelerated, this time to the energy of 450 GeV. They are then transferred to the LHC via one of the two transfer lines. After the filling of both counter rotating beams with bunch-trains, the protons are accelerated while the magnetic field in the bending dipoles is simultaneously rising until the collision energy is reached. The beams are then «squeezed» to have as many collisions as possible in the collision points of ATLAS and CMS. For LHCb, the beams are slightly defocussed to reduce the instantaneous luminosity (see later). If no technical problems occur, the beams can stay several hours in the LHC until they are finally dumped and the cycle starts again.

---

<sup>3</sup>A quench is a sudden loss of superconductivity in a magnet.

<sup>4</sup>The idea is that the proton bunches sit around the zero voltage point of the electrical field with a negative slope. If a particle arrives too early, it will see a positive electrical field and will be put into a longer orbit, which will lead to a delay in the next arrival. The opposite happens for a particle arriving too late: it will see a negative electrical field and therefore arrive earlier in the next arrival.

<sup>5</sup>Not only protons can be brought into collision at the LHC but also heavy ions. These are mainly analysed by the ALICE experiment.

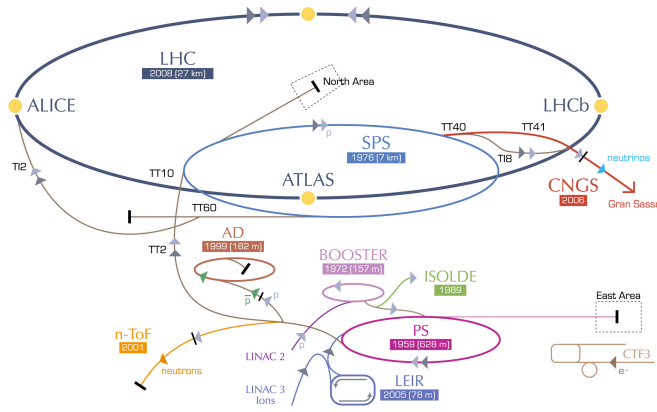


Figure 2.2.: Accelerator chain at CERN. For the proton proton collisions at the LHC, LINAC2, BOOSTER, PS and SPS are of importance. Not to scale. Figure taken from Ref. [46].

### 2.2.2. LHC performance

The LHC has two main performance numbers: Energy per beam and luminosity. The luminosity is a measure of how many collisions happen when two bunches of particles collide with each other. It is linked to the number of collisions per second and the cross-section:

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma, \quad (2.1)$$

where  $\frac{dN}{dt}$  is the number of collisions per second,  $\sigma$  the cross-section of the process in question and  $\mathcal{L}$  the (instantaneous) luminosity. To have a measure for the total amount of data acquired, one often uses the integrated luminosity  $\mathcal{L} = \int \mathcal{L} dt$ , which then has the units of an inverse area. For a Gaussian beam distribution, the luminosity can also be written as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \varepsilon_n \beta^*} F, \quad (2.2)$$

where  $N_b$  is the number of particles per bunch,  $n_b$  the number of bunches per beam,  $f_{rev}$  the revolution frequency,  $\gamma_r$  the relativistic gamma factor,  $\varepsilon_n$  the normalised transverse beam emittance,  $\beta^*$  the beta function at the collision point and  $F$  the geometrical luminosity reduction factor due to the crossing angle at the interaction point. It takes into account that the beams don't collide head on. The emittance is a measure of the distribution of the particles in space and momentum and  $\beta^*$  is a measure of how compressed the beam is at the collision point. All of these quantities are functions of accelerator properties

and can (in theory) be calculated from first principles — they do not rely on the underlying physics processes.

The achieved values for these quantities in 2011 and the design values of the LHC are listed in Table 2.1.

Table 2.1.: Key numbers of the LHC, reached in 2011 and in the design. Note that for some numbers the reached values exceed the designed ones.

Quantity	2011 (peak)	Design
Energy	3.5 TeV	7 TeV
Luminosity in ATLAS/CMS/LHCb	$3.6/3.6/0.4 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$	$10/10/0.2 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$
Number of bunches	1380	2808
Number of protons in bunch	$1.49 \cdot 10^{11}$	$1.15 \cdot 10^{11}$
Stored energy	115 MJ	362 MJ
Time between two collisions	50 ns	25 ns

## 2.3. The LHCb experiment

The LHCb experiment [47] [48] [49], located at Point 8 of the LHC, is a single-arm spectrometer with a forward geometry, covering angles from about 10 mrad to approximately 300 (250) mrad in the horizontal (vertical) direction. It is shown in Fig. 2.4. Its forward design arises from the fact that  $b$  and  $\bar{b}$  quarks are produced in pairs and predominantly in the forward or backward direction. This geometry allows LHCb to reconstruct a large fraction of the produced particles containing a  $b$  or  $\bar{b}$  quark while only covering a small solid angle (see Fig. 2.5).

The coordinate system of LHCb is oriented such that the positive  $z$ -direction points from the interaction point to the muon system along the beam-pipe, the  $y$ -coordinate is vertical starting from the interaction point to the surface and perpendicular to the LHC ring<sup>6</sup> and the  $x$ -coordinate is in the so-called «bending plane»<sup>7</sup>, such that all three axes form a right-handed coordinate system. The dimensions of LHCb are roughly  $20 \text{ m} \times 10 \text{ m} \times 10 \text{ m}$  and the detector is situated in the same cavern as DELPHI, which was a (magnificent)  $4\pi$  detector. To make the maximum use of the available space in the already existing cavern, the collision point had to be displaced to allow the most efficient construction of a forward spectrometer.

<sup>6</sup>The negative  $y$ -axis does not point towards the center of gravity, due to the inclination of the LHC as mentioned in Sect. 2.2.1.

<sup>7</sup>Charged particles are mainly bent inside this plane by the magnetic field.

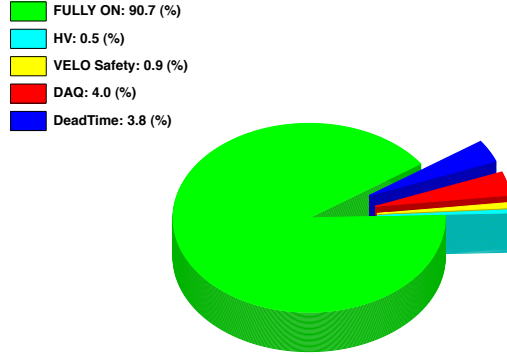


Figure 2.3.: Efficiency of the data taking of LHCb in 2011 with the different contributions to the inefficiency. «HV» stands for high voltage and «DAQ» for data acquisition.

The LHCb detector was designed for an average instantaneous luminosity of  $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ , with a peak luminosity of  $5 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  at a centre-of-mass energy  $\sqrt{s} = 14 \text{ TeV}$ . The average number of inelastic proton-proton collisions ( $\mu$ ) would have been about 0.4 per bunch-crossing to achieve the nominal integrated luminosity of  $2 \text{ fb}^{-1}$  per year. However, in 2010,  $\mu$  was set to values up to 2.7, as there were fewer bunches in the beam than the LHC was designed for. The total integrated luminosity was about  $38 \text{ pb}^{-1}$ . In 2011, the average  $\mu$  was 1.5 with 1380 bunches per beam at maximum. The total integrated luminosity<sup>8</sup> was about  $1.2 \text{ fb}^{-1}$ . It should be noted that not all collisions in the LHCb detector can be used later for data analysis as there are times when certain detector parts are not operative. An overview of the contributors to this inefficiency is shown in Fig. 2.3.

LHCb has decided against using the full luminosity provided by the LHC (as CMS and ATLAS do) for several reasons. The forward region is dominated by a very high flux of particles which creates high occupancies in the detectors and induces radiation damage. Running at full luminosity would have put severe constraints on the choice of detector materials and segmentations to use. Moreover, separation of primary and secondary vertices is crucial for many analyses in LHCb, a task which is more difficult with large pile-up. Additionally, the ability to reconstruct all tracks of importance in the event degrades with increasing number of interactions. The instantaneous luminosity for LHCb was therefore chosen as a compromise of all these parameters.

<sup>8</sup>It is worth mentioning that measuring the luminosity in a hadron collider is not a trivial task. The best precision achieved in 2010 was 3.5% [50].

The LHCb detector can be grouped in three parts. The track reconstruction system aims at determining the three components of the particles' momenta, the particle identification system at determining the particle types. These two properties completely describe each individual particle and therefore the full event. Finally the trigger system selects the events of interest for physics analyses.

- **Track reconstruction:** The track reconstruction systems comprise a silicon microstrip detector close to the interaction point, called the VELO, which measures the position of the primary vertices and the impact parameters (IP) of the track with high precision. A further silicon microstrip detector (TT) is placed before the dipole magnet. Its task is to improve the momentum resolution of reconstructed tracks and reject pairs of tracks that in reality belong to the same particle. The magnetic field of the dipole magnet bends the flight path of the particles in the  $x$ - $z$  plane and therefore allows the determination of their momenta by comparison of the slopes before and after the magnet. The tracking system is completed by the T stations, which, together with the information from the VELO, determine the momentum and flight direction of the particles. The T stations use different technologies for detecting particles: silicon microstrips close to the beam pipe and straw-tubes in the outer regions. All tracking detectors are characterized by having a high spatial resolution (in one or two coordinates) and a low material budget.
- **Particle identification:** The particle identification systems use different physical principles for their task. The two Ring Imaging Cherenkov Detectors, RICH1 and RICH2, use the fact that particles emit light under a characteristic angle when traversing certain materials, an effect that allows for the distinction between different types of hadrons. The electromagnetic and hadronic calorimeters, ECAL and HCAL, measure the energy of the impinging particles by fully absorbing them. They are assisted by two smaller subdetectors in front of them, the SPD and the PS, that allow resolving ambiguities in the identification. The muon system, which is placed at the most remote position within LHCb seen from the interaction point, is built to identify muons. Its stations (M1 to M5) register particles that traverse the detector and the iron shields between the muon stations unaffected, a property of charged particles that only muons exhibit.
- **Trigger:** The information from every collision is too high for all of it to be stored. Moreover, many of the collisions are not of interest for physics analyses. Therefore LHCb has a three stage trigger system to reduce the amount of data collected to a rate which can be written to disk. The first level, called L0, is hardware based and directly implemented in the detector. The second and third stage, HLT1 and HLT2, are software based and execute algorithms that (partially) reconstruct the event and then decide if they are of further interest or not.

The detector components are described in detail below.

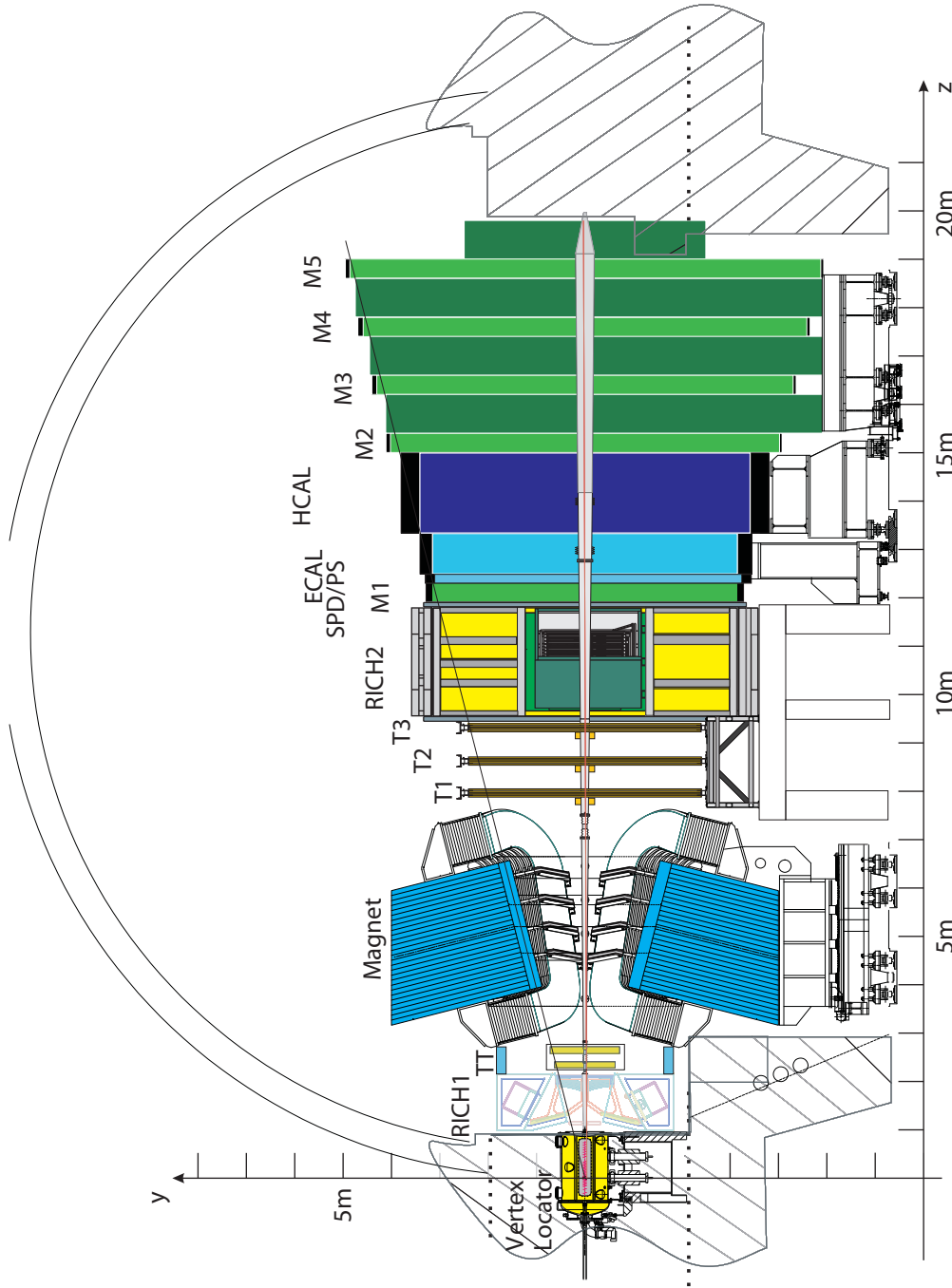


Figure 2.4.: The LHCb detector in all its glory! Figure taken from Ref. [51].

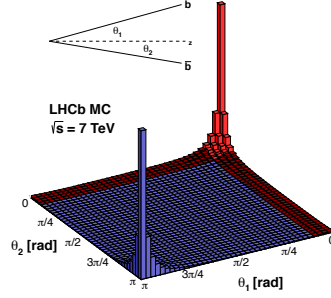


Figure 2.5.: Distribution of  $b$  and  $\bar{b}$  quarks in simulated proton-proton collisions at  $\sqrt{s} = 7$  TeV, with  $\theta$  being the angle between the quark and the beam axis. Red is the LHCb acceptance. The correlation and peak in the forward / backward direction is clearly visible. Figure taken from Ref. [52].

### 2.3.1. The magnet

LHCb features a warm dipole magnet with two aluminum coils inside an iron yoke [53] [54] [55]. The magnet has an integrated field of  $\int \vec{B} d\vec{l} = 4$  Tm, where the main field component is along the  $y$ -axis. Its shape and field strength are illustrated in Fig. 2.6. Note that there is practically no field at the position of the VELO, while there is only a small field at the position of the TT and the T stations. The magnetic field polarity can be reversed, a procedure which is regularly undertaken during operation of the detector to minimise systematic effects.

In order to track the particles through the LHCb detector, a precise knowledge of the magnetic field map is of great importance. The magnetic field was measured with Hall probes, whose results could then be compared to magnetic field simulations. More information on this can be found in Ref. [56].

### 2.3.2. The Vertex Locator

The Vertex Locator (VELO) [58] [48] [49] is the subdetector closest to the proton-proton interactions. It is designed to precisely measure the position of the primary vertex, which is crucial for many analyses. It also provides a first measurement of the particles' flight path. It is therefore essential for the active material of the VELO to be as close to the interaction region as possible. Furthermore, two space dimensions need to be measured. The VELO uses silicon microstrip sensors in a  $r$ - $\phi$  geometry and consists in total of 21 «stations» of silicon sensors. Each stations has a detector module on the left and on the right side of the beam axis, where each module comprises a  $r$ - and  $\phi$ -measuring sensor. The  $r$ - $\phi$  geometry, in contrast to a rectilinear one, was chosen to simplify the fast reconstruction of tracks and



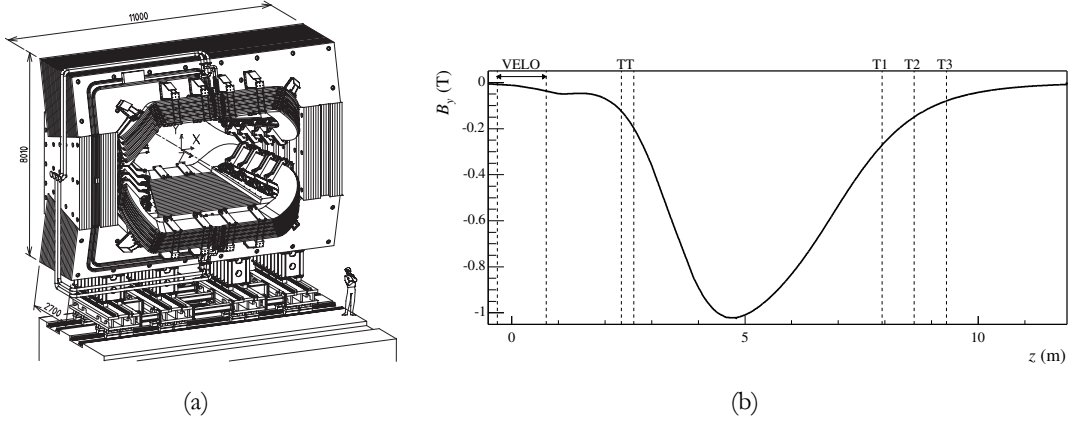


Figure 2.6.: Technical drawing of the dipole magnet in LHCb (a) and strength of the  $y$ -direction of the magnetic field as a function of the  $z$ -position (b). Figures taken from Ref. [49] and Ref. [57].

vertices. Two additional stations were designed for vetoing pile-up events. They are located upstream of the nominal interaction point and have  $r$ -sensors only.

The VELO covers a pseudorapidity-range  $\eta$  between 1.6 and 4.9 and is able to reconstruct tracks emerging from a vertex with  $|z| < 10.6$  cm, see Fig. 2.7. The requirement that a track should at least cross three VELO stations defines the position of the three most downstream stations to be at approximately  $z = 65$  cm. The full length of the VELO is about 1m.

A VELO module has a diameter of about 90 mm and covers a bit more than  $180^\circ$  in azimuth, where a small hole is left free in the middle for the colliding beams, see Fig. 2.8. The sensors consist of a silicon  $n$ -bulk with implants of  $n^+$  forming the strips and  $p$ -spray to isolate them from each other. This design is very radiation hard, superior to more commonly used  $p^+$ -on- $n$  type silicon. For the  $r$ -sensors, the pitch for the strips closest to the beam is  $40 \mu\text{m}$ , which increases linearly to  $101.6 \mu\text{m}$  for the outermost strip. This design makes sure that strips closer to the interaction point get a larger weight for the impact parameter determination. Each strip is divided into four sectors covering  $45^\circ$  to reduce the occupancy and the capacitance. The  $\phi$ -sensors are divided into two regions: the inner stops at a radius of 17.25 mm and has a pitch of  $78.3 \mu\text{m}$ , the outer one starts at 17.25 mm and has a pitch of  $39.3 \mu\text{m}$ . A naive design with a single strip running radially would have too large occupancies and a very high pitch at the outer edge. Furthermore, the strips of the inner and outer regions do not join as straight lines but have a kink at the boundary. The modules are then placed such that adjacent  $\phi$ -sensors have the opposite skew, which corresponds to a stereo view allowing the distinction between real hits and hits which do not belong to a particle having traversed the VELO. Both,  $r$ - and  $\phi$ -sensors

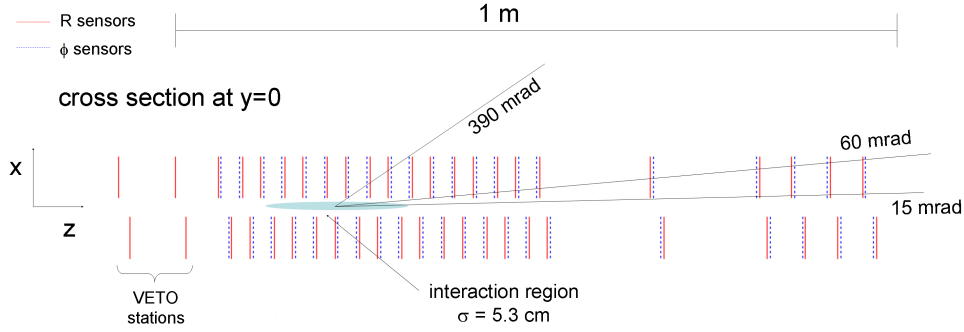


Figure 2.7.: Diagram showing the spacing of VELO modules along the  $z$ -axis and the respective arrangement of  $\phi$ - and  $r$ -sensors. Figure taken from Ref. [49].

are  $300\text{ }\mu\text{m}$  thick. An illustration of both sensor types is given in Fig. 2.8.

The modules for each detector half are placed in an aluminum-walled box, which is under vacuum. The side facing the beam is realised as a RF foil, made out of an aluminum alloy, to suppress RF pickup from the LHC beams. Furthermore, the beam vacuum and the vacuum inside the VELO-box are separated by the RF foil to protect the beam vacuum from outgassing of VELO sensors. The RF-foil represents a considerable amount of the total material budget of the VELO ( $\approx 40\%$ ) and has a notable impact on the track reconstruction efficiency for particles crossing it (see Chap. 4).

The proton beams are not constant in shape when going from the injection-state to the collision-state. The modules therefore have to be retractable to avoid the danger of being damaged. The parking position places the two halves about 6 cm away from each other; when the LHC is in proton-proton collision, the halves are moved in until the sensors have a distance from the beam of only 8 mm. A beam interlock system is in place which inhibits moving the VELO close to the beam when the beam conditions are not stable.

The hit resolution of an  $r$ -sensor of the VELO is shown in Fig. 2.9 for two projected angles of the track. Note that the best resolution is  $4\text{ }\mu\text{m}$ , which is the best vertex detector resolution achieved at the LHC.

### 2.3.3. The Tracker Turicensis

The Tracker Turicensis (TT) [49] [59] [60] is located between RICH1 and the magnet and consists of four planes of silicon microstrip sensors with a pitch of  $183\text{ }\mu\text{m}$  and a strip length up to 37 cm. The sensors have a thickness of  $500\text{ }\mu\text{m}$  and use single-sided  $p^+$ -on- $n$  silicon. They are arranged in half modules, which consist of seven silicon sensors which are read out in sectors, containing four and three sensors (away from the beam pipe) or four, two and one sensors (close to the beam pipe). An illustration for the last type can be seen

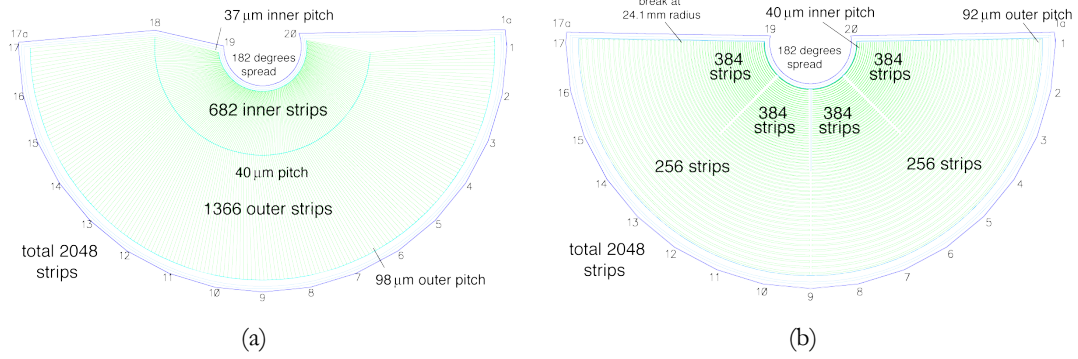


Figure 2.8.: Sketch for the  $\phi$ - (a) and  $r$ - (b) sensors of the VELO. Note the kinks in the strips of the  $r$ -sensors (to reduce the occupancy) and the increasing strip pitches when moving from the inside to the outside. Figures taken from Ref. [58].

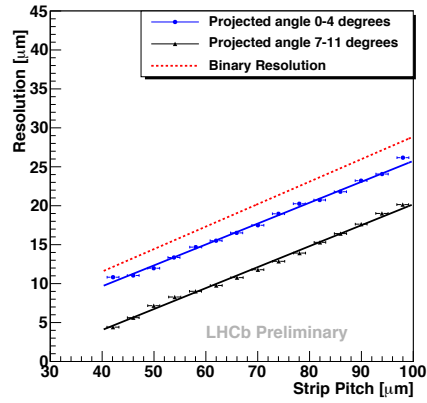


Figure 2.9.: Hit resolution of the VELO for two different angles of the track as a function of the strip pitch, in comparison to the pure binary resolution. The best achievable resolution is about 4  $\mu\text{m}$ .

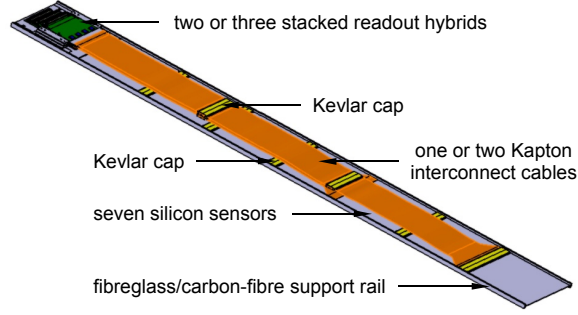


Figure 2.10.: Half module of the T1 with one, two and four sectors bonded together for the readout. Figure taken from Ref. [61].

in Fig. 2.10. The different length of these read-out sectors is due to the different amount of particles traversing the detector. The particle flux is highest in the strips closest to the beam pipe, whilst it falls off by about two orders of magnitude in the outermost regions of the detector. The readout electronics are placed at the upper or lower edge of the half module, lying outside the acceptance of LHCb. Two half modules are joined together to span the full height of the T1 except for the part above and below the beam-pipe. While the first and the last layer have sensors with strips running vertically<sup>9</sup>, while the second and third layer are tilted with an angle of  $-5^\circ$  and  $+5^\circ$  with respect to the vertical axis. An illustration of the first and the second T1 layer can be seen in Fig. 2.11. This setup allows for a stereo view and a tilted layer is consequently called a «stereo layer». The first and second layer are called T1aX and T1aU, while the third and fourth layer are referred to as T1bV and T1bX. T1a and T1b denote the stations with the first two and the last two layers. The two stations are separated by about 27cm, while the layers inside T1a or T1b are much closer to each other. Note that adjacent modules within a layer are staggered by about 1 cm in the  $z$ -direction and have overlap regions in  $x$  of a few millimeters. This avoids gaps in the acceptance and simplifies the relative alignment of the individual modules. Each layer in T1a has 30 half-modules while in T1b the layers have 34 half-modules, to cover the larger area required for the full acceptance at a larger value of  $z$ . In total, the silicon forms an active detector area of  $8.4\text{m}^2$  with a total of 143'360 readout channels.

T1 is housed in two half-stations (one for every side of the beam pipe) which can be retracted horizontally for maintenance. During operation, the inside of the box forms one large volume. The modules are mounted on cooling plates at the top and bottom of the box, which provide support for the modules and cool the electronics. The box volume is continuously flushed with nitrogen to avoid condensation on the cold surfaces, moreover the temperature inside the box is kept at about  $8^\circ\text{C}$ .

<sup>9</sup>Vertical with respect to the LHC coordinate system.

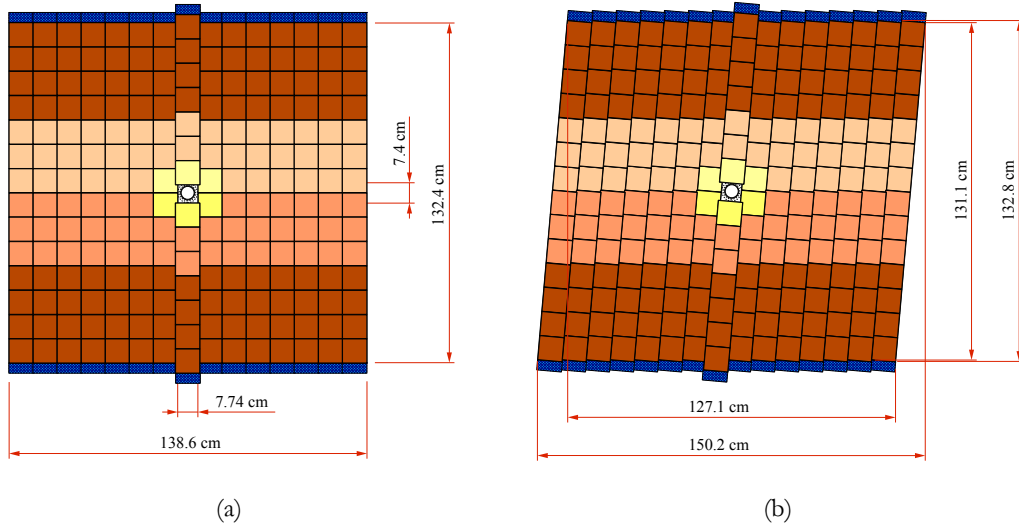


Figure 2.11.: Sketch of the first layer in TT (TTaX) in (a) and the second layer in TT (TTaU) with a tilt of  $5^\circ$  of the modules in (b). The different colours correspond to different readout sectors, the blue rim is the position of the readout electronics. Figures taken from Ref. [61].

Originally TT was built to be an integral part of the trigger system (its name was then «Trigger Tracker»), however it was never used in this role. It is mainly used to improve the momentum resolution for particles traversing the whole detector or reconstructing flight paths of particles decaying outside the VELO. It also improves the rejection of ghost tracks, which are tracks that do not belong to a real particle having transversed the detector. The hit resolution is about  $50 \mu\text{m}$  with an occupancy in the range 3.5% (close to beam pipe) to 0.35% (outermost sector).

### 2.3.4. The Inner Tracker

The Inner Tracker (IT) [49] [62] forms the inner part of the tracking stations (T stations) closest to the beam pipe. As the particle flux is the highest in these regions, silicon sensors were chosen as detecting devices. The IT consists of three stations, where each station comprises four detector boxes which are arranged around the beam pipe, see Fig. 2.12. Each box houses four detector layers and each layer has a total number of seven modules. As in the TT, the first and last layer within a box run vertically, while the second and third layer are tilted by  $-5^\circ$  and  $+5^\circ$  with respect to the vertical, respectively. The modules to the left and right of the beam pipe consist of two sensors, while the ones on top and bottom of the beam pipe only have one sensor. The IT (and all other detectors downstream of the

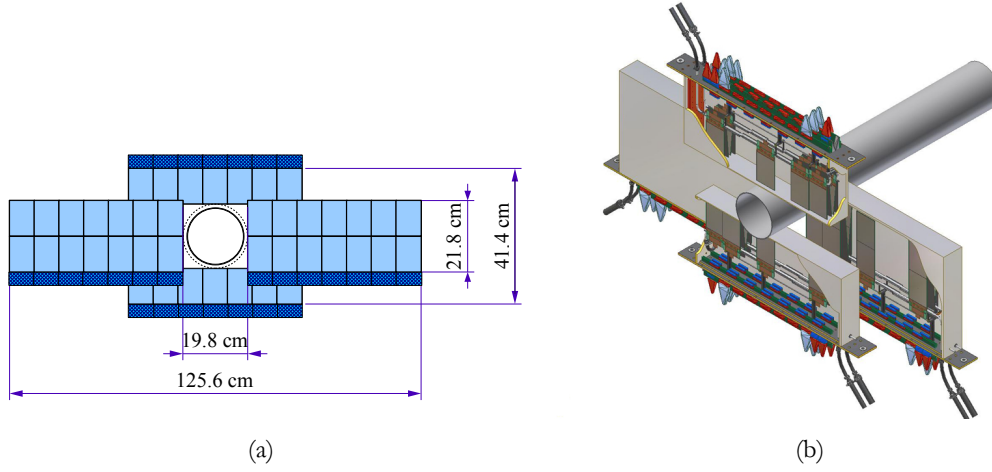


Figure 2.12.: Sketch of a layer in IT (a) with the light blue being the silicon sensors and the dark blue being the position of the readout electronics. (b) shows the four boxes of an IT station with its four layers of modules. Figures taken from Ref. [49].

magnet) are built with the detector layers running vertically with respect to the center of gravity, not the LHC coordinate system. This choice was taken to simplify the construction of the support structures for the large and heavy detectors downstream of the magnet.

The silicon sensors used in the IT slightly differ from the ones used in the TT. They are single-sided  $p^+$ -on- $n$  sensors with a size of  $7.6 \text{ cm} \times 11 \text{ cm}$  and a thickness of  $320 \mu\text{m}$  for one-sensor modules and  $410 \mu\text{m}$  for two-sensor modules. They have a strip pitch of  $198 \mu\text{m}$  and 384 readout strips per sensors, which leads to a similar resolution as in the TT. Similarly to the situation in the TT, the electronics are cooled via a cover plate which incorporates cooling pipes, through which the cooling agent  $\text{C}_6\text{F}_{14}$  is pumped. The box is then also flushed with nitrogen to avoid condensation on the cold surfaces. The whole IT is mounted on large support frames which are attached to rails outside the detector acceptance on the top and the bottom. As the readout electronics and cooling system cannot be placed outside the acceptance due to the position of IT, the material budget for the Inner Tracker is non-uniform: Close to the beam it adds up to  $0.035 X_0$ <sup>10</sup> per station while for the narrow region of the cooling pipes it can add up to  $0.3 X_0$  per station.

The occupancy in IT ranges from about 2.5% (strips closest to the beam) to about 0.5% (outermost strip) for the boxes left and right of the beam pipe while they are in the range of 0.3% to 0.5% in the boxes above and below the beam pipe.

<sup>10</sup>The concept of  $X_0$ , called the radiation length, is explained in Sect. 3.2.1.

### 2.3.5. The Outer Tracker

The outer part of the T stations, called Outer Tracker (OT) [63] [49] is built as an array of straw-tube modules. These drift-tubes are arranged in two staggered, displaced layers to form a module and have an inner diameter of 4.9 mm. A sketch of a module is shown in Fig. 2.14. The modules are arranged in three stations with four layers each, with again the middle two layers in a station tilted by  $\pm 5^\circ$  with respect to the vertical, similar to the arrangement in the IT. Two types of modules exist: Long ones, which have a length of about 5 m and contain 256 straw-tubes, and short ones with about half the length of the long ones and half the number of straw-tubes. They are used above and below the IT. All together, this results in about 55'000 channels for the whole OT. Except for the cross-shaped region occupied by the IT, the OT covers the full LHCb acceptance of 250 (300) mrad in the non-bending (bending) plane which leads to a total area of  $5971 \times 4850 \text{ mm}^2$ . The shape and extent of the region assigned to IT was chosen so that the occupancy in the OT does not exceed 10% for an instantaneous luminosity of  $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ . The arrangement of the IT and the OT is shown in Fig. 2.13.

A single drift-tube has a cylindrical shape and its volume is filled with a gas-mixture of  $\text{Ar}/\text{CO}_2/\text{O}_2$ , which allows for a short drift time. The drift-tubes are electrically split into a top and bottom half with a separate readout for both halves. This split is not done at the same vertical position of the two monolayers in a module to avoid insensitive regions, furthermore there is no split for the short type of modules. Again, the detector is separated into a left and right unit, which both can be retracted. The design of the OT allows placing all electronics and support structures outside the acceptance (contrary to the situation for the IT). However, the straw-tubes still contribute a significant amount to the material budget: one station (*i.e.* 8 monolayers of drift-tubes) add up to about 3.2% of  $X_0$  which is significantly more than for TT or IT.

The distance of closest approach of the particle to the anode wire is used to measure the coordinate of the particle. It can be determined with the knowledge of the relation of the distance of closest approach and the drift time, which is the experimentally measured quantity. This then allows for a spatial resolution of about  $200 \text{ } \mu\text{m}$  for a single cell with a hit efficiency of more than 99%.

### 2.3.6. The Ring Imaging Cherenkov Detectors

LHCb uses two Ring Imaging Cherenkov Detectors (RICH) [49] [64] [48] for particle identification, mainly for pion / kaon separation. Their working principle is based on the Cherenkov effect [65]: particles traversing a medium (called radiator) faster than the speed of light in this medium will emit photons at an angle  $\theta_c$ . Knowing the refractive index of the material being traversed, this angle can then directly be related to the speed of the particle:

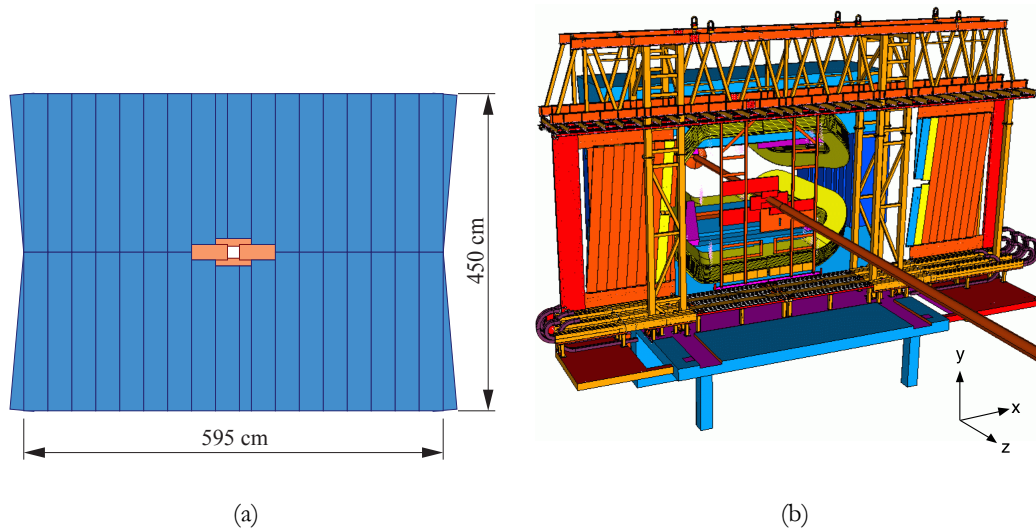


Figure 2.13.: Sketch of one of the T stations with the OT in blue and the IT in orange (a) and situation of one T station with infrastructure (b). The OT is retracted while the IT is at its nominal position. Figures taken from Refs. [62] and [49].

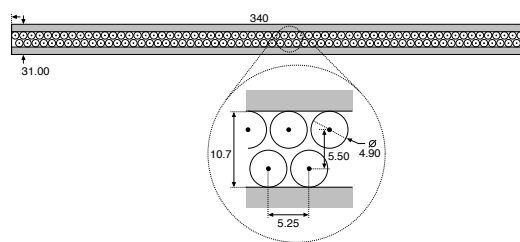


Figure 2.14.: Cross section of an OT straw-tube module with the two monolayers displaced with respect to each other. Figure taken from Ref. [49].



$$\cos \theta_c = \frac{1}{n\beta}, \quad (2.3)$$

with  $n$  the refractive index and  $\beta = \frac{v}{c}$  the (normalised) velocity of the particle. With the knowledge of the particle's momentum, the invariant mass and therefore the particle type can be determined.

The first Cherenkov detector, RICH1, is situated between the VELO and the TT and covers the full LHCb acceptance. A sketch is shown in Fig. 2.15. It uses two kinds of radiators: plates of aerogel with a refractive index of 1.03 at  $\lambda = 400$  nm to provide separation for particles with a momentum of a few GeV/ $c$ ; and C<sub>4</sub>F<sub>10</sub> as a gaseous radiator filling the volume of RICH1. The C<sub>4</sub>F<sub>10</sub> has a refractive index of 1.0014 at  $\lambda = 400$  nm and provides kaon-pion separation up to 60 GeV/ $c$ . The Cherenkov light is then reflected via a spherical and a plane mirror into an array of Pixel Hybrid Photon Detectors (HPDs) at the top and bottom of the RICH1 (outside the acceptance), which amplify the signal. These HPDs accelerate the photoelectron, which results from the Cherenkov photon impinging on the surface of the HPD, onto a pixel sensor. This sensors provides a position measurement allowing the reconstruction of the «rings» created by the Cherenkov light. The radius of these rings is proportional to the Chereknov angle. The photon detectors are sensitive to stray magnetic fields, hence they have to be embedded into a magnetic shield structure (so called «shield boxes»), which need to attenuate the external field by a factor of 20 without having an impact on the overall field in this region.

The second Cherenkov detector, RICH2 is located between the T stations and the first muon station. It has a reduced angular acceptance of 15 mrad to 120 (100) mrad in the bending (non-bending) plane. It uses CF<sub>4</sub> as a radiator, with a refractive index  $n$  of 1.0005 at  $\lambda = 400$  nm and it is used for particle identification for high-momentum particles from about 15 GeV/ $c$  up to 100 GeV/ $c$ . Unlike RICH1, RICH2 has the HPDs and mirrors situated to the left and right of the beam pipe, incorporating a magnetic shield structure similar to the one in RICH1.

The Cherenkov angles for the three different radiators used in RICH1 and RICH2 are shown in Fig. 2.15, while the discrimination power between pions and kaons is shown in Fig. 2.16.

### 2.3.7. The calorimetry system

The calorimetry system [66] [49] serves multiple purposes. On the one hand it is used to reconstruct photons and particularly  $\pi^0$  with good precision, on the other hand it is part of the particle identification system providing information about the energy and position of photons, electrons and hadrons. Furthermore it is part of the hardware trigger, making fast decisions after the interaction.

The calorimeters consist of three parts: A scintillating pad / preshower detector (SPD/PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The SPD/PS

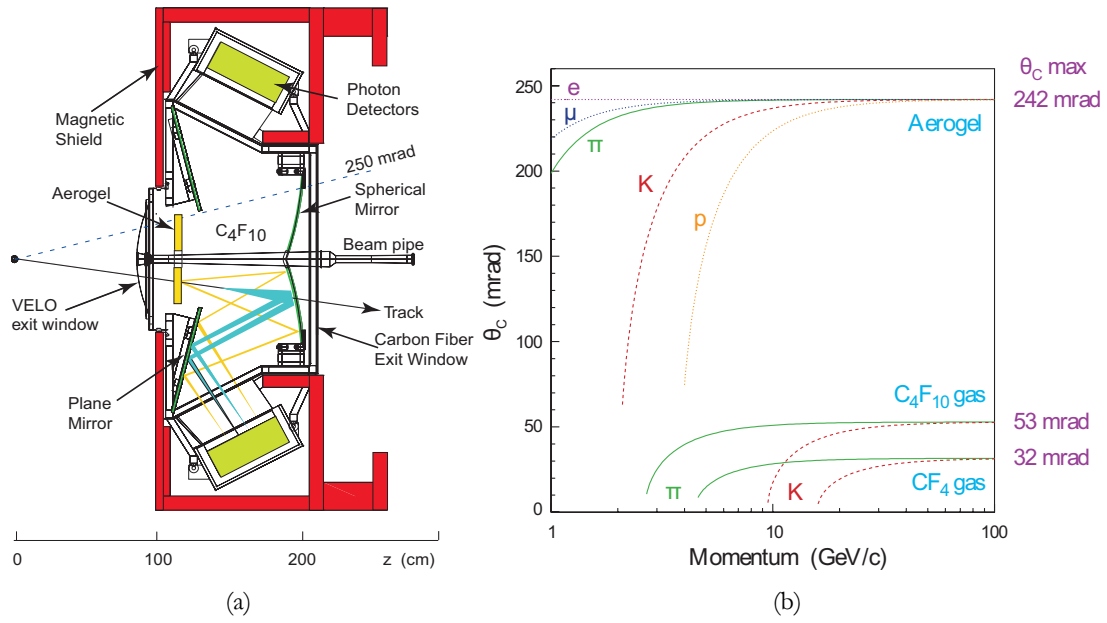


Figure 2.15.: Sketch of RICH1 (a) with light paths for photons emerging in the aerogel and C<sub>4</sub>F<sub>10</sub>. In (b) the Cherenkov angle for the different radiators in different momentum regimes is shown. Both figures taken from Ref. [49].

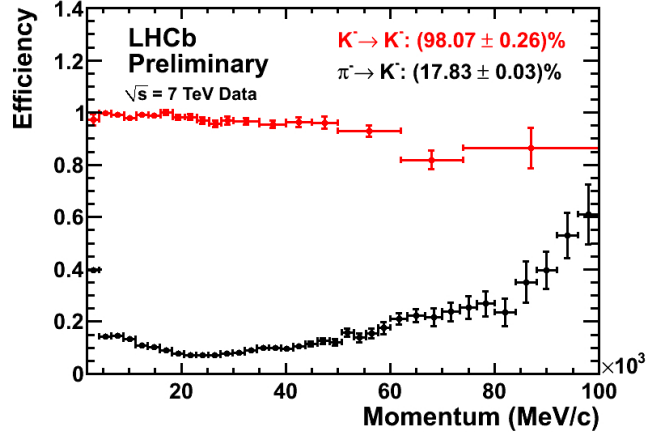


Figure 2.16.: The discrimination power between kaons and pions as a function of the momentum. In red are kaons correctly identified as kaons and in black are pions misidentified as kaons. The two samples are separated using a cut on the combined likelihood  $DLL_{K\pi}$ , distinguishing hypotheses which are more kaon- or more pion-like. The  $DLL_{K\pi}$  variable uses mostly information from the RICH detectors.

and the ECAL follow the same substructure by dividing the active region in three parts, with a more coarse granulation when moving away from the beam pipe. The HCAL has a division into two subparts. The segmentations are illustrated in Fig. 2.17.

The SPD/PS consist of a 15 mm thick lead absorber, which is sandwiched by plates of scintillation pads. The pads are read out via wavelength shifting fibres guiding the light to photomultipliers. The first layer of scintillating pads is used to record deposited ionisation and therefore reject  $\pi^0$  with high  $E_T$ . The lead absorber is used to start the showering, which can then be detected in the PS. It provides a longitudinal segmentation to the ECAL which is used to improve the discrimination between pions and electrons, as they have different shower lengths. The SPD and PS pads have dimensions of approximately  $4 \times 4 \text{ cm}^2$ ,  $6 \times 6 \text{ cm}^2$  and  $12 \times 12 \text{ cm}^2$ , depending on the region.

The ECAL is built as a «shashlik» calorimeter *i.e.* an alternating structure with absorbing layers made of lead (2 mm thick) and detecting layers made of scintillating material (4 mm thick, polystyrene). In total, there are 66 Pb and scintillating layers that form a stack of 42 cm length, which corresponds to  $25X_0$ . The readout of the scintillating tiles is similar to the one in the SPD/PS. The outer acceptance boundary of the ECAL matches the one from the tracking stations (300/250 mrad), the inner acceptance boundary (25 mrad) is limited by the radiation dose level close to the beam pipe. The resolution of the ECAL is approximately [49]:

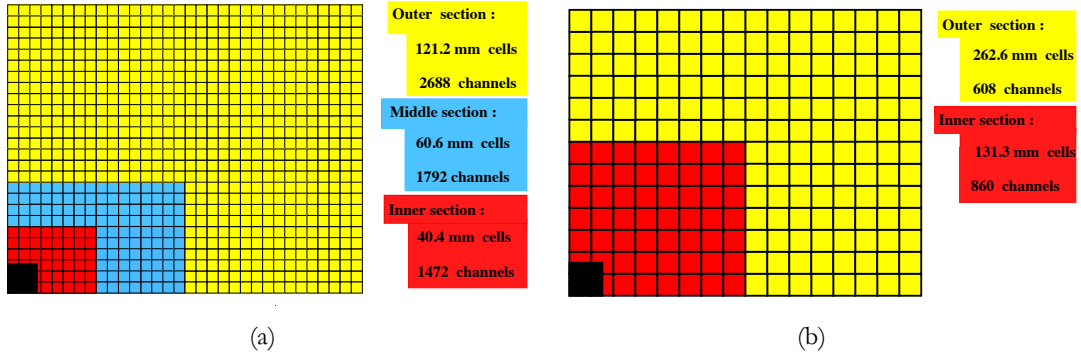


Figure 2.17.: Layout of the SPD/PS/ECAL (a), showing the three segmentations, and layout of the HCAL (b) with only two segmentations. Figures taken from Ref. [66].

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\text{GeV}}} \oplus 1\%. \quad (2.4)$$

The first term is the resolution due to statistical fluctuations and the second term is due to the readout.

The HCAL is built as a sampling device as well, made of iron as an absorber and scintillating pads as active material. However, in the HCAL the scintillating tiles and the iron absorbers run parallel to the beam axis. In total three tiles and three absorbers define the length of the HCAL, corresponding to  $5.6\lambda_{int}$ <sup>11</sup>. The scintillating material is again read out as in the SPD/PS. The resolution of the HCAL is about [49]:

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E/\text{GeV}}} \oplus 10\%, \quad (2.5)$$

where the two terms have the same definition as for the ECAL.

### 2.3.8. The muon system

The muon system in LHCb [67] [68] [69] [49] consists of five stations: The first one (M1) is placed before the calorimeters, the other four stations (M2 to M5) are situated behind the HCAL and separated from each other with layers of 80 cm of iron, which serves as an absorber for all particles not being muons or neutrinos (see Fig. 2.18). A muon needs to have a momentum of about 6 GeV/c to reach all muon stations. The muon system covers an acceptance between 20 (16) mrad and 306 (258) mrad in the bending (non-bending)

<sup>11</sup>The concept of the hadronic interaction length  $\lambda_{int}$  is introduced in Sect. 3.2.2.



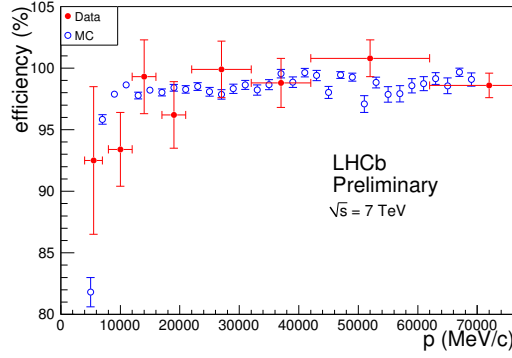


Figure 2.19.: Efficiency of the muon system in 2010 collision data (red) and simulation (blue) for identifying real muons using the «isMuonLoose» classifier (explained in Sect. 6.1.2).

a muon standalone track for triggering. The resolution also decreases when moving away from the beam pipe, as for larger angles multiple scattering is dominant.

In total, the muon system uses 1368 MWPCs with a gas mixture of Ar, CO<sub>2</sub> and CF<sub>4</sub>. In M2 to M5, a MWPC is made of four «gas gaps» *i.e.* having four volumes for gas and four anode wires, see Fig. 2.18. This improves the redundancy in case of a failure of a single gap. Two adjacent gaps are coupled together in an OR-logic. In M1, the MWPCs only have two gaps to reduce the amount of material in front of the calorimeters, which adds up to  $0.26X_0$  for the first muon station. The triple-GEM chambers (12 in total) consist of two superimposed triple-GEM detectors each with an active area of  $20 \times 24 \text{ cm}^2$ , which are combined in a logical-OR as well.

The efficiency of the muon chambers was tested on 2010 collision data and it is around 99% for all chambers [71]. It has also been checked that the loss of one gap does not cause a significant loss of efficiency in M2 to M5. The efficiency to identify a real muon as a muon by a loose classifier called «isMuonLoose» is shown in Fig. 2.19. The classifier is explained in Sect. 6.1.2.

### 2.3.9. The trigger system

Although it is not a visible part of the LHCb detector, the trigger is a crucial component of the experiment. It allows the quick rejection of events which are not of interest in order to bring the amount of data to a rate which can be stored on disk. The trigger system consists of two parts: L0, the first stage of the trigger implemented on hardware-level; and the HLT, a two-stage software trigger executed on event-filter farms, to confirm the L0 decisions and to perform a (partial) event reconstruction for further data reduction. The basic working principle of a trigger system is that thresholds are defined for different properties and that

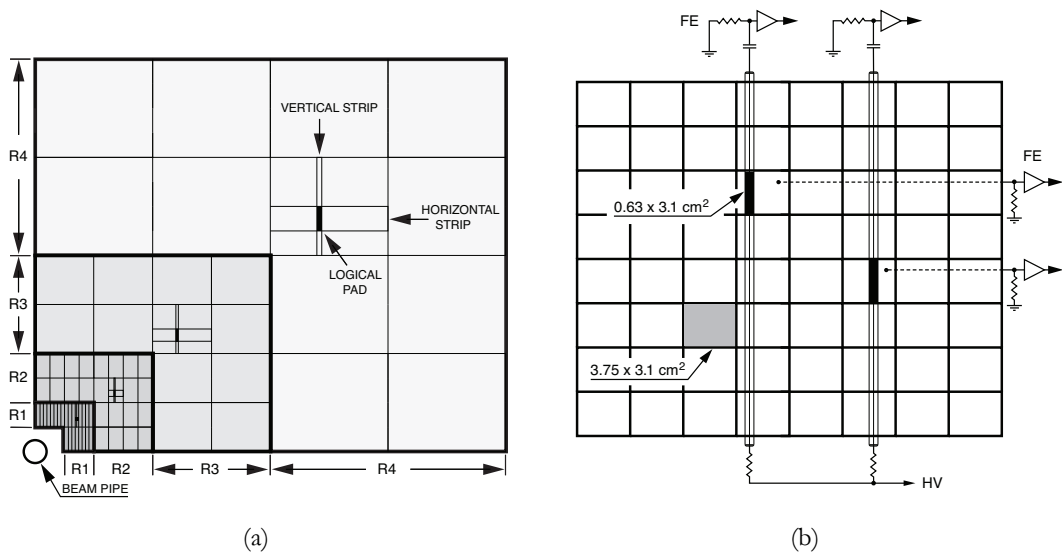


Figure 2.20.: (a) shows the segmentation of the muon system in regions and the definition of a logical pad. (b) shows how the wires and pads are read out by the front-ends (FE) and how a logical pad is obtained by a logical AND of both, vertical and horizontal, signals. Figures taken from Ref. [49].

Table 2.2.: Segmentation of the muon system for different stations and regions. The numbers correspond to the size of the logical pads in  $x$  and  $y$ , all dimensions are given in cm. See Fig. 2.20 for an illustration of the different sizes of the regions.

	<b>M1</b>	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
<b>R1</b>	$1 \times 2.5$	$0.63 \times 3.1$	$0.67 \times 3.4$	$2.9 \times 3.6$	$3.1 \times 3.9$
<b>R2</b>	$2 \times 5$	$1.25 \times 6.3$	$1.35 \times 6.8$	$5.8 \times 7.3$	$6.2 \times 7.7$
<b>R3</b>	$4 \times 10$	$2.5 \times 12.5$	$2.7 \times 13.5$	$11.6 \times 14.5$	$12.4 \times 15.5$
<b>R4</b>	$8 \times 20$	$5 \times 25$	$5.4 \times 27$	$23.1 \times 29$	$24.8 \times 30.9$

the event is accepted if one of these properties exceeds the threshold.

The L0 [72] [49] trigger is run synchronously with the LHC clock and reduces the rate of visible pp collisions<sup>12</sup> to about 1.1 MHz, which is the maximum rate at which the full detector can be read out. Three systems contribute to the L0 trigger: the VELO-pile up detector, the calorimeters and the muon system. The pile-up detector was designed to reject events with multiple pp collisions. However, as these types of events were desired for the collision data taken in 2010 and 2011 it was only used for the selection of beam-gas events. The calorimeters (SPD, PS, ECAL and HCAL) are used to select high  $E_T$ <sup>13</sup> particles (photons, electrons, hadrons) and assign a particle hypothesis to them. The candidate with the highest  $E_T$  for every species is selected and triggers the event if the  $E_T$  is above a given threshold. Furthermore the SPD is used to provide an estimate for the charged particle multiplicity in the event. A cut on the SPD-multiplicity was set in 2010 and 2011 to reject too busy events which would take too much time to process and have a too high occupancy in the OT. The muon system performs a standalone reconstruction of muon tracks by forming straight tracks out of hits in all five stations and calculates their transverse momentum under the assumption that the muon originated at the interaction point (more on this strategy will be discussed in Chap. 4). There is a single muon trigger and a dimuon trigger which have different thresholds.

HLT1 was designed to confirm the output of L0 by partially reconstructing the event around the object that triggered L0, while using information from different subdetectors. This task is for example accomplished by muon triggers which combine the information from the VELO with the one from the muon stations [73]. Other triggers exist to confirm information from the calorimeters. Additionally, a different type of HLT1 trigger exists, which does not rely on information from the L0 and is based on a single track with high momentum, large impact parameter and good track quality, as for example the daughter

<sup>12</sup>A visible pp collision is required to have at least two tracks reconstructed in LHCb.

<sup>13</sup> $E_T$  stands for «transverse energy» and is defined as  $E_T = E \cdot \sin \theta$  with  $\theta$  the angle between the beam axis and the particle momentum.



particles of  $B$  decays will exhibit [74]. In total, HLT1 has an output rate of approximately 50 kHz.

The output rate of HLT1 is low enough to allow HLT2 to perform a reconstruction of the event which is very close to the offline reconstruction. A great variety of trigger lines exist in HLT2, ranging from lines selecting prompt and detached muons [73] to «topological lines», selecting n-body  $B$  decays starting from two charged daughter tracks of the  $B$  [75]. Furthermore, several exclusive lines exist, *e.g.* for the decays  $B_s^0 \rightarrow \phi\gamma$  and  $B^0 \rightarrow K^{*0}\gamma$  [76]. In 2011, the output rate of HLT2 was up to 3 kHz.



## 3. Track reconstruction in LHCb

In this chapter, an introduction to track reconstruction in LHCb is given. First, the basic principles of the interaction of particles with matter are explained. Second, the tracking strategy in LHCb and its particularities are described in detail. Finally, pattern recognition algorithms are explained with the focus on the algorithm which adds information of the  $TT$  station to tracks formed out of particles traversing the whole LHCb tracking system.

### 3.1. The anatomy of a flight path

Tracking is the procedure to reconstruct the flight path of a charged particle. It is essential in (almost) every particle physics experiment to determine the momentum and the flight direction of the particles.

Track reconstruction has seen diverse experimental methods: from photographing the tracks in cloud- or bubble chambers or even reconstructing the path with wire models in early studies of cosmic rays [77] to electronic readout from silicon detectors in today's particle detectors at the LHC. In LHCb (as in most modern particle physics experiments) the tracking detectors are sampling detectors [78]. The detector volume is interspersed with devices measuring the position of the impinging particles in this device, see Chap. 2. Knowing the position of the particle at different points in space and the map of the magnetic field (see Sect. 2.3.1) allows the reconstruction software to determine the full flight path.

It is worth mentioning that there is no standard pattern recognition method that can be applied to any experiment. The specific layout of each detector leads to specific algorithms which cannot easily be compared with each other. Furthermore, no «software package» exists which could easily be implemented into the framework of the detector software to perform track reconstruction.

### 3.2. Passage of particles through matter

When particles pass through matter, they interact with the material. Two distinct types of interactions exist: electromagnetic and hadronic interactions.<sup>1</sup>

---

<sup>1</sup>In principal also weak interactions should be taken into account. However, this is without relevance for tracking.

### 3.2.1. Electromagnetic interactions

One can distinguish several types of electromagnetic interaction of charged particles<sup>2</sup>.

**Ionisation** Ionisation occurs when the particle traveling through the material excites an atom to a higher state or ionises it completely via interactions with its outer electrons. The mean loss of energy is described by the Bethe-Bloch-Formula [8]:

$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \log \left( \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (3.1)$$

where  $K = \frac{4\pi e^4}{c^2 m_e} N_A$ , with  $e$  the electron charge,  $c$  the speed of light,  $m_e$  the mass of the electron,  $N_A$  the Avogadro number,  $z$  the charge of the particle (in electron charges),  $Z$  the atomic number of the absorber,  $A$  the atomic mass of the absorber,  $\beta = \frac{v}{c}$ ,  $I$  the mean excitation energy (in eV),  $T_{\max}$  the maximum kinetic energy transferred to a free electron in a single collision and  $\delta(\beta\gamma)$  an energy correction due to polarisation.

The Bethe-Bloch-Formula describes the mean energy loss for velocities  $0.1 < \beta\gamma < 1000$  with a precision of a few percent for intermediate- $Z$  materials. For lower  $\beta\gamma$  values corrections have to be introduced in Eq. 3.1 to account for the finite speed of the electrons in the atomic shells. For higher values, the energy loss due to radiative processes becomes dominant (see later).

For a qualitative discussion, the Bethe-Bloch-Formula can be approximated to [80]:

$$-\left\langle \frac{dE}{dx} \right\rangle \approx \rho \left( 2 \text{ MeV} \frac{\text{cm}^2}{g} \right) \frac{Z^2}{\beta^2}. \quad (3.2)$$

This equation shows the most important dependencies of energy loss due to ionisation: the speed of the particle, its electric charge and the density of the material. An illustration of the energy loss for different particles and the Bethe-Bloch curve is shown in Fig. 3.1.

**$\delta$ -rays**  $\delta$ -rays are electrons which are knocked out of atoms of the material by the incoming (high energy) particle. These electrons travel a finite distance and can themselves ionise further atoms. Energetic  $\delta$ -rays are rare and are not considered any further in the context of track reconstruction.

**Coulomb scattering** Also called Rutherford Scattering, this type of interaction occurs between the particle and the nuclei of the atoms in the material. The interaction does not lead to an energy loss, but changes the trajectory of the particle. Many subsequent scatterings are then termed «multiple scattering», which will be discussed later.

---

<sup>2</sup>In addition to the interactions mentioned here, also Cherenkov radiation (see Sect. 2.3.6) and transition radiation [79] exist. However, these effects are not important in the context of track reconstruction in LHCb.

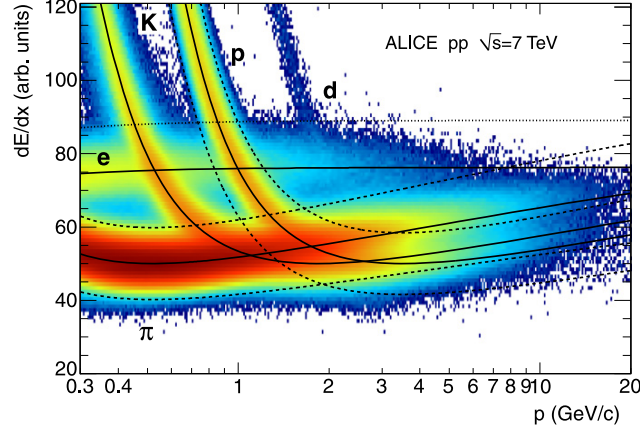


Figure 3.1.: Specific energy loss in the time projection chamber of the Alice experiment as a function of momentum with superimposed Bethe-Bloch lines for various particle species. The dashed lines show the pion and proton exclusion bands. The dotted line corresponds to a cut for electrons. Figure taken from Ref. [81].

**Bremsstrahlung** Bremsstrahlung occurs when a charged particle emits a photon under the influence of the field of the atomic nucleus. It is the dominant source of energy loss for electrons in a high-energy physics collider experiment<sup>3</sup>. The concept of radiation length,  $X_0$ , is useful to quantify the energy loss due to Bremsstrahlung where  $X_0$  is the average distance an electron has to fly in a given material to reduce its energy by a factor  $e$ . The energy after the distance  $x$ , where  $x$  is the distance of material traversed, is given by:

$$E_e(x) = E_0 \cdot e^{-\frac{x}{X_0}}. \quad (3.3)$$

### 3.2.2. Hadronic interactions

In addition to electromagnetic interactions, hadrons (charged and neutral) can also undergo hadronic interactions. Thereby they break up atomic nuclei, which leads to the release of protons and neutrons (called spallation), or they undergo deep inelastic scattering, which produces new hadrons, mostly pions. The quantity of interest is the nuclear interaction length,  $\lambda_{int}$ , which describes the mean free path of a particle between two hadronic interactions. In (gaseous) hydrogen, for example,  $\lambda_{int}$  is  $6.21 \cdot 10^5$  cm while in lead it is 17.59 cm [82]. The initial particle is often lost in hadronic interactions and cannot be tracked anymore. The cross section depends on the type of particle, its charge and momentum. The total amount of material up to the last tracking station is about 20% of a

<sup>3</sup>The critical energy for which Bremsstrahlung dominates for electrons is approximately given by  $E_c = \frac{800 \text{ MeV}}{Z+1.2}$  [80]. For other particles with mass  $M$ , Bremsstrahlung is suppressed by a factor  $\left(\frac{m_e}{M}\right)^2$ .

hadronic interaction length. This topic is discussed again in Sect. 4.8.3.

### 3.2.3. Multiple Scattering

Multiple scattering is a sequence of Coulomb scatterings that lead to a change of the flight direction of the particle. It is described in detail with the Bethe-Molière formalism [83] for all angles. For small scattering angles, the distribution of the projected angle can be approximated by a Gaussian distribution with a width  $\theta_0$  [84] given by:

$$\theta_0 = \frac{13.6 \text{ MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left( 1 + 0.038 \log \frac{x}{X_0} \right) \quad (3.4)$$

where  $\beta c$ ,  $p$  and  $z$  denote the velocity, the momentum and the charge number of the incoming particle respectively and  $x$  is the thickness of the material. This formula is accurate to better than 11% for all materials.

## 3.3. Representation of the flight path of a particle

Tracking consists of two separate parts: Pattern recognition and fitting. The pattern recognition bundles together all the hits a single particle left in all the subdetectors. A hit in the context of tracking is a well defined point or line in space which corresponds to a part of a tracking detector which yielded an electronic signal (not identified as noise). The idea is to create a set of hits for each particle that has travelled through the active region of the detector. These sets should ideally be disjoint between different particles and not neglect any possible hits or add fake ones.<sup>4</sup> The track fit then tries to find the best description of the flight path of the particle and its momentum, including the corresponding errors. The track fit in LHCb is done using the Kalman formalism [85] [86] [87] which will be explained below.

### 3.3.1. Parametrisation of a track

A track can be parametrised by line segments which are tangential to the flight path of the particle for any given point in space. These line segments and their respective covariance matrices (to represent the errors) are called states and can be described by 5 parameters

---

<sup>4</sup>In reality this is of course not possible, see the discussion on efficiency and purity of the TT hit adding algorithm in Sect. 3.6.3.

(and a 5 dimensional matrix) at a given  $z$ -position:

$$\vec{x}_z = \begin{pmatrix} x \\ y \\ t_x \\ t_y \\ \frac{q}{p} \end{pmatrix} \text{ with } t_x = \frac{\partial x}{\partial z}, t_y = \frac{\partial y}{\partial z} \quad (3.5)$$

where  $q$  is the charge of the particle and  $p$  its momentum. States can be anywhere along the flight path but are normally chosen at the measurement planes of the tracking detectors. Two operations with track states are of particular interest: propagation and projection. Propagation provides the track state at a position  $z_2$  given the track state at a position  $z_1$ , following the formula:

$$\vec{x}_{z_2} = f_{z_1 \rightarrow z_2}(\vec{x}_{z_1}), \quad (3.6)$$

where  $f$  denotes the track propagation function. Multiple scattering is taken into account in this step by enlarging the error of the propagation when the particle travels through a piece of material. The propagation function can be highly non-linear in regions with a magnetic field<sup>5</sup>. A simple matrix form is used in the absence of a magnetic field:

$$f_{z_1 \rightarrow z_2}(\vec{x}_1) = F_{z_1 \rightarrow z_2} \vec{x}_{z_1} = \begin{pmatrix} 1 & 0 & z_2 - z_1 & 0 & 0 \\ 0 & 1 & 0 & z_2 - z_1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \vec{x}_{z_1}. \quad (3.7)$$

The projection step describes the relation between a measurement and a state via the formula:

$$m_z = h_z(\vec{x}_z), \quad (3.8)$$

where  $h_z$  is the projection function. Again, this relation simplifies if a measurement directly provides one of the coordinates of the track state. In this case, the projection relation becomes linear and  $h_z$  can be written as a matrix.

### 3.3.2. Kalman filtering for the track fit

Track fitting is a procedure to provide the best estimate of the parameters which describe the particle's flight path, such as momentum or impact parameter. The Kalman filter technique is specially suited for a track fit in a high energy physics experiment due to the following reasons:

- Possible exclusion of outliers in the pattern recognition in the prediction step.

---

<sup>5</sup>In this case, an adaptive 5<sup>th</sup> order Runge-Kutta method is used for propagation, see Ref. [88].

- **Computational speed:** No inversions of large matrices are needed (contrary to a least-squares-estimation).
- Multiple scattering and energy loss can be taken into account very naturally.

While these are important advantages for an efficient tracking algorithm, the final result is equivalent to a least-squares fit. The Kalman fit can be divided into three steps (the calculations can be found in Ref. [57]):

- **Prediction:** This step predicts the parameters of a state at the position  $z_2$  given a state at the position  $z_1$ , using Eq. 3.6. In order to do this, it needs at least one initial state,  $\vec{x}_0$ , which has to be provided by an initial fit in a track finding algorithm. This step takes multiple scattering into account by increasing the errors (*i.e.* modification of the covariance matrix).
- **Filtering:** The filtering step updates the state at  $z_2$  with information of the measurement at this position.

These two steps are repeated until all measurements, selected by the pattern recognition, are added to the track. At this point only the track state at the position of the measurement added last incorporates the full information.

- **Smoothing:** In this step all other track states are updated in the reverse direction using the «smoother equations» to provide the best estimates of the track states at all given positions.

An illustration of the Kalman formalism can be seen in Fig. 3.2.  
In LHCb the track fit is run from the T-stations to the VELO.

## 3.4. Tracks types in LHCb

There are five distinct (standard) track types in LHCb<sup>6</sup>, categorised by the individual sub-detectors (and combinations of them) involved in the reconstruction of the tracks.

- **Long tracks:** Long tracks are the most important tracks in the LHCb experiment. They traverse the detector starting in the VELO up to the T stations. Long tracks leave hits in the VELO and the T-stations, but not necessarily in TT. They are made with two different algorithms called «Forward Tracking» and «Track Matching» (see Sect. 3.5.1 and 3.5.2).

---

<sup>6</sup>For special purposes, there exist further track types. Some of them will be discussed in Chap. 4.



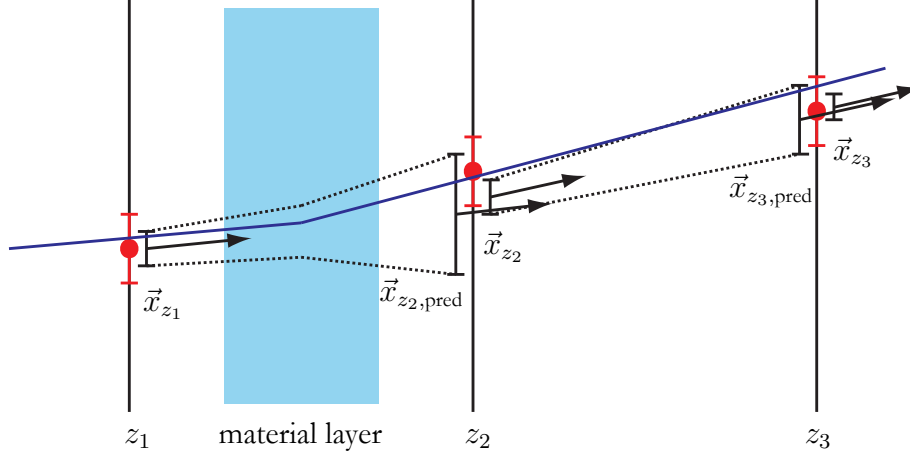


Figure 3.2.: Illustration for the Kalman formalism.  $\vec{x}_{z_i}$  denote the filtered states,  $\vec{x}_{z_{i,\text{pred}}}$  the predicted states, the red dots correspond to the measurements. Note that the error is enlarged due to the presence of a material layer which could lead to multiple scattering. The Kalman filter corrects for the change in the direction of the trajectory by pulling the filtered states towards the measurement.

- **Downstream tracks:** Downstream tracks are made by charged, long living particles with a origin vertex considerably displaced from the interaction point. Typical examples are  $K_s^0$  mesons or  $\Lambda$  baryons. They only leave (sufficient) hits in the TT and the T-stations. Due to the lack of hits in the VELO and therefore the longer lever arm, the momentum resolution is worse for these types of tracks than for long tracks.
- **Upstream tracks:** Upstream tracks are formed by low momentum particles whose flight paths get strongly deflected by the magnetic field so that they don't reach the T-stations anymore. They are used in the RICH1 reconstruction.
- **T tracks:** T tracks are tracks that could only be reconstructed in the T stations. The corresponding particles may come from material interactions or decays of very long living particles. They are used in the RICH2 reconstruction and can be used for the internal alignment of the T-stations.
- **VELO tracks:** VELO tracks have hits in both the  $r$ - and  $\phi$ -sensors of the VELO, but cannot be matched to hits in the other tracking stations. They are used for reconstructing primary vertices.

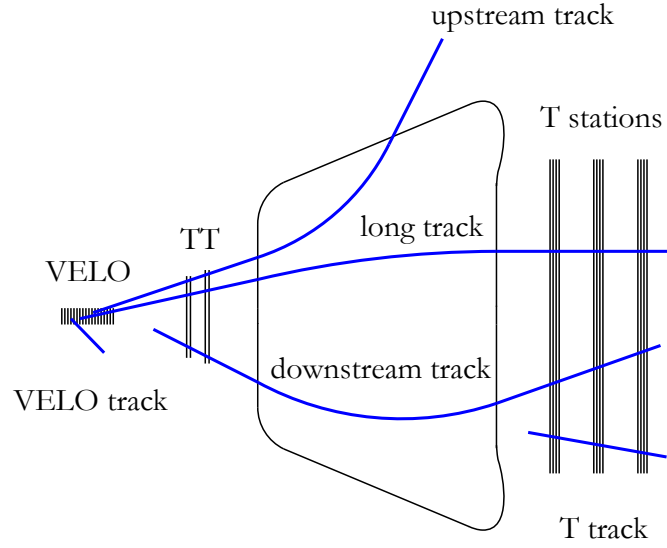


Figure 3.3.: Sketch of the LHCb detector in x-z-projection with VELO, TT, the magnetic dipole and the T-stations. The blue lines correspond to the five (standard) track types.

## 3.5. Pattern recognition algorithms

There are two algorithms in place in LHCb to find long tracks: «Forward Tracking» and «Track Matching». As long tracks are the most important track types for physics analyses, they will be briefly discussed.

### 3.5.1. Forward Tracking

The forward tracking [89] [90] is based on an optical analogy: The magnetic field can be seen as a lens which kicks the flight path of a particle from one direction into another at a well defined plane in space, much like a lens can be seen as giving a kink to a light ray in geometrical optics. Although the magnetic field occupies a finite volume in space and therefore bends the flight path continuously, the final result in the tracking stations outside the magnetic field is the same. The only complication to this simple model is the fringe field of the magnet, which will lead to a non-constant position of the «magnet bending» plane. However, as this change in position is rather small (order of cm), the analogy can still be exploited. As seen in Sect. 3.3, a track has five parameters: two for the position, two for the slope and one for the momentum. If the precise shape of the magnetic field is known and multiple scattering is neglected, the Forward Tracking allows the full reconstruction of the flight path.

The procedure starts with a «seed»<sup>7</sup> in the VELO. These tracks are straight as there is no magnetic field in the VELO. They determine the position and the slope. To fix the last parameter (the momentum), a single point in the T-stations is needed<sup>8</sup>. The algorithm therefore combines a VELO track with every hit in a x-plane in the T-stations to reconstruct a (possible) flight path. It then projects all measurements along this flight path in the T-stations onto a plane at a fixed z-position. This transformation is called a «Hough transform», see Fig. 3.4. The measurements corresponding to the particle of the VELO seed should cluster, while other measurements should follow a random distribution. In the following, the hits in the cluster are fitted and outliers are removed using a  $\chi^2$  criterium. In this stage also hits from the stereo layers are searched for clusters, fitted and again selected using  $\chi^2$ . Finally, the best track candidate is selected based on multiple track quality criteria. It should be noted that TT hits have not been added up to now. This will be done in a separate step and explained later in Sec. 3.6. Also note that a long track does not necessarily need to have TT hits.

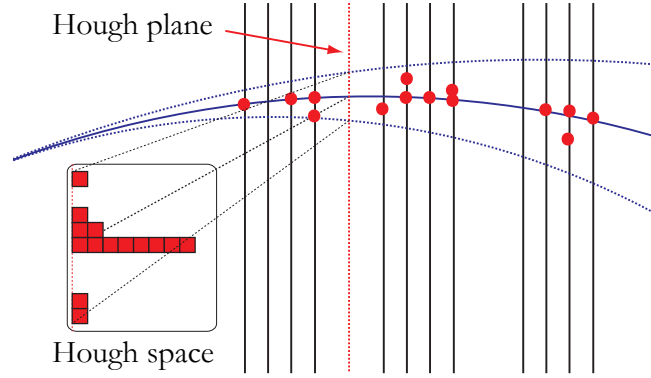


Figure 3.4.: Sketch of a Hough transform, with extrapolated track (solid, blue) and search windows (dashed, blue). The hits are projected along the extrapolated track onto the Hough space. Hits belonging to the same track cluster around the position of the extrapolated track while other hits are randomly distributed.

### 3.5.2. Track Matching

A different method to reconstruct particles that traverse the full detector is called «Track Matching» [91] [92]. In a first step it combines all standalone tracks from the T-stations

<sup>7</sup>A seed is the starting point for every track and is constructed by collecting hits using very general principals like physically possible angles. Often seeds are constructed in parts of the detector without a magnetic field. An example can be found in Ref. [78].

<sup>8</sup>The bending of the flight path only occurs in the  $x$ -direction to first order. Therefore a single measurement providing the  $x$ -coordinate is enough.

(found by an algorithm called «PatSeeding» [93] [94]) and from the VELO (made with an algorithm called «FastVelo» [95]). Both seeds of a VELO-T station pair are extrapolated to the magnet bending plane to calculate their distance in  $x$  and to the end of the T stations to calculate their distance in  $y$  (practically no bending occurs in the  $y$ -direction). This distance as well as the difference in the slopes is then used to form a compatibility criterion on this VELO-T station-track pair, which is used as a measure for the «goodness» of the track.<sup>9</sup> As in the Forward Tracking, TT hits are only added after these steps are completed.

### 3.5.3. Additional pattern recognition algorithms

**VELO stand-alone tracking:** Due to the absence of a magnetic field in the VELO, the pattern recognition in the vertex locator is looking for straight tracks. First, the measurements in the  $r$ -sensors are used to find tracks in the  $r$ - $z$ -projection while the measurements in the  $\phi$ -sensors are added afterwards. The tracks are fitted using a  $\chi^2$  minimization. More information can be found in Refs. [96] and [95].

**T station stand-alone tracking:** The T stations are affected by the fringe field of the magnet; a standalone tracking with a (limited) momentum estimate is therefore possible. The algorithm is described in Refs. [93] and [94]. Tracks are reconstructed by combining a hit in every station into a parabola, trying all reasonable three-hit-combinations. A search region around this parabola is defined, where more hits are added to the track. In the next step, the hits from the stereo layers are added to the track candidate. As tracks can traverse both IT and OT, a special procedure has to be applied to find these tracks.

**Downstream tracking:** Downstream tracks have hits in the TT and the T-stations. The pattern recognition [97] is similar to the forward tracking, except that a T-station track is extrapolated through the magnet into the TT, where corresponding hits are selected (using a Hough transform) and added to the track. This task is somewhat difficult as the four layers of the TT don't leave much room for redundancy.

**Upstream tracking:** To form upstream tracks, VELO tracks are extrapolated into the TT where hits are selected and added similarly to the procedure in the downstream tracking [98]. One main difference is the presence of a small magnetic field between the VELO and the TT which makes the momentum estimate more difficult.

---

<sup>9</sup>Note that an earlier implementation of the matching algorithm (called «TrackMatching», where the name refers here to the name of the algorithm and not the method) used Kalman-fitted T-station tracks extrapolated to  $z = 830$  mm, where they got matched with linearly extrapolated VELO-tracks. A «matching- $\chi^2$ » was used as a quality measure. While having a similar performance in terms of finding the right tracks, it was much slower due to the Kalman-filter-step and was not used anymore in the 2011 data taking. The algorithm which was used in the 2011 data taking is called «PatMatch».

### 3.5.4. Final track selection

As there are two independent algorithms to reconstruct long tracks and several track types form a subset of other types (*e.g.* downstream tracks are essentially shortened long tracks), a mechanism must be in place to disentangle them. On the one hand this is done by only using hits which have not been used in a previous stage of the track reconstruction, *e.g.* only VELO tracks that could not be made into long tracks are used for upstream tracks. On the other hand, both Forward Tracking and Track Matching reconstruct long tracks. For this purpose a «Clone Killer» [99] is in place. This is an algorithm which compares two tracks based on the number of measurements. If they share many hits, the track with the smaller amount of associated hits is discarded.

## 3.6. Adding TT hits to long tracks

In both long track finding algorithms, TT hits are only added after a VELO T-station track has been formed. Originally there were two algorithms to add TT hits, «PatAddTTCoord» for the Forward Tracking and «AddTTClusterTool» for the Track Matching. For the 2010 and 2011 data taking «PatAddTTCoord» was used exclusively. The original algorithm is described in Ref. [90] with a more detailed description given below<sup>10</sup>. The exact form of the parametrisations will be explained in Sect. 3.6.2.

### 3.6.1. The TT hit adding algorithm

The algorithm «PatAddTTCoord» consists of a series of selection cuts and clustering/projection procedures. The most important dependence of these cuts is the momentum, as the precision of the predicted position largely depends on this quantity. This stems from the fact that multiple scattering decreases with increasing momentum as seen in Eq. 3.4.

The following steps are executed for every VELO-T station track in the event.

***y* compatibility:** In a first step all hits in the TT in a given event are considered. For every possible hit, the track state closest to the middle of TT is then extrapolated linearly in the *y*-coordinate to the *z*-position of the hit to estimate the predicted *y*-position in the measurement plane. The hit has to fulfill the following equation:

$$|y_{\text{pred}} - \frac{1}{2}(y_{\text{beg}} + y_{\text{end}})| < \frac{1}{2}(y_{\text{beg}} - y_{\text{end}}) + y_{\text{To1}}, \quad (3.9)$$

where  $y_{\text{pred}}$  denotes the predicted *y*-position at the *z*-position of the hit and  $y_{\text{beg}}$  and  $y_{\text{end}}$  are the *y*-positions of the beginning and end of a silicon strip and  $y_{\text{To1}}$  is a tolerance parameter. As TT is a strip tracker, a hit can be seen as a line with a start and end point. The equation ensures that only hits compatible with the *y*-position of the predicted state (within a given tolerance) are considered further.

---

<sup>10</sup>The algorithm «AddTTClusterTool» is described in Refs. [91] and [57].

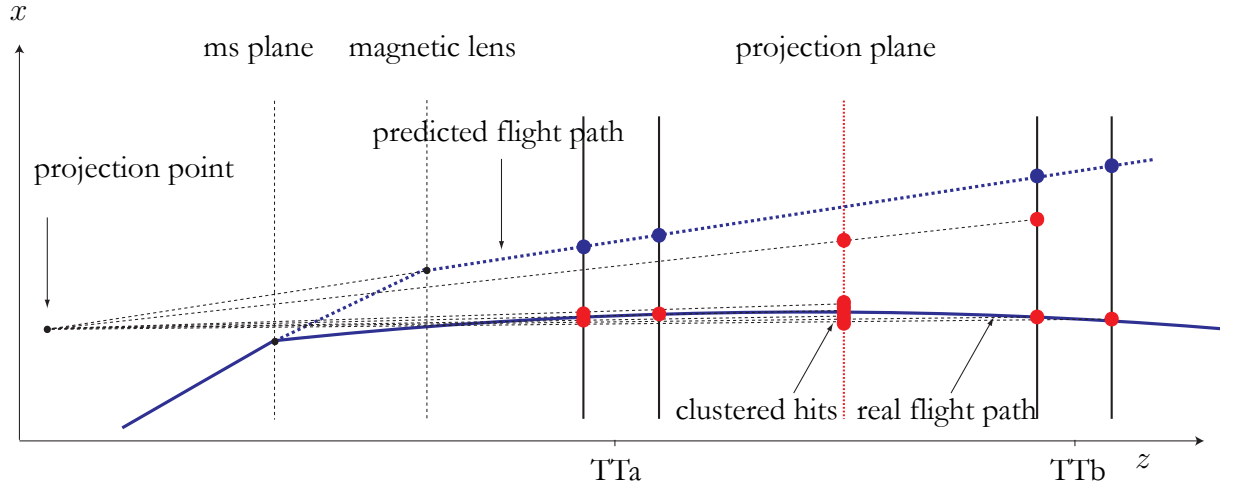


Figure 3.5.: Scheme of the track extrapolation in the  $x$ - $z$ -plane in TT. The predicted flight path (blue dashed line) has a kink in the magnetic lens to take the bending of the track into account. The real flight path (blue solid line), in addition to its bending, has an additional kink due to multiple scattering («ms planes»). The hits are projected into a plane between TTa and TTb («projection plane») with a fixed vanishing point («projection point») to take this into account. The projected hits belonging to the real flight then cluster at a certain distance from the predicted flight path in the projection plane, while other hits will be more randomly distributed.

**$x$ -coordinate prediction** Now also the  $x$ -position is predicted in TT. For this, the state closest to the middle of TT is extrapolated to the  $z$ -position of the hit using a simplified model for the magnetic field. As only a stray field is present, an «effective» magnetic plane can be defined using the same optical model as in the Forward Tracking. The trajectory of the particle can then again be approximated with two straight lines which have a kink at the magnetic plane as shown in Fig. 3.5. This yields the following prediction:

$$x_{\text{pred}} = x_{\text{state in VELO}} + (z_{\text{hit}} - z_{\text{state in VELO}}) \cdot \left. \frac{dx}{dz} \right|_{\text{VELO}} + \mathbf{tt}_{\text{par}} \cdot \frac{q}{p} \cdot (z_{\text{hit}} - z_{\text{mag. field}}), \quad (3.10)$$

where  $\mathbf{tt}_{\text{par}}$  is a slope parameter which has to be tuned in simulation and  $\left. \frac{dx}{dz} \right|_{\text{VELO}}$  is the  $x$ -slope of the track in the VELO.

The hit in question has to fulfill  $|x_{\text{pred}} - x_{\text{hit}}| < \mathbf{x}_{\text{To1}}$  where  $\mathbf{x}_{\text{To1}}$  is a tolerance parameter which depends on the momentum of the particle.  $x_{\text{hit}}$  is the  $x$ -position of the hit and is the  $x$ -position in the middle of the strip in case the hit is from a stereo layer.

If there are less than three hits fulfilling this condition, they will not be added and the long track will not have TT hits associated to it. If there are three or more hits, they are processed further.

**Projection**  $x_{\text{pred}}$  and  $x_{\text{hit}}$  are then projected onto a plane between TTa and TTb to take possible multiple scattering into account. The vanishing point of the projection is called «projection point». It reproduces the effect of a scattering plane for particles between the VELO and TT, see Fig. 3.5. Note that the position of the projection point is not equal to the position of the multiple scattering plane as the flight paths of particles are bent. The idea is that hits belonging to the track will cluster at a certain distance from the predicted position in the projection plane while other hits will have a more random distribution. The formula for the projection is:

$$x_{\text{hit, pred; proj}} = x_{\text{hit, pred}} \cdot \frac{z_{\text{TT proj. plane}} - z_{\text{proj. point}}}{z_{\text{hit, pred}} - z_{\text{proj. point}}} \quad (3.11)$$

$$\Rightarrow \text{proj} = x_{\text{pred; proj}} - x_{\text{hit; proj}}, \quad (3.12)$$

where the difference between the projections of the hit position and the predicted position is just called «projection» for simplicity.

**Forming groups** The hits are now grouped together. The first group starts with the hit which has the smallest projection. Further hits are added as long as they fulfill the equation:

$$\text{proj}_{\text{hit}} - \text{proj}_{\text{first hit}} < \sqrt{\text{majAxProj}^2 \cdot \left(1 - \frac{\text{proj}_{\text{first hit}}}{\text{minAxProj}}\right)^2}, \quad (3.13)$$

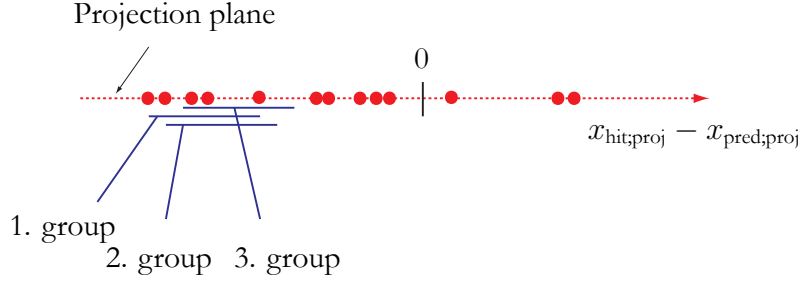


Figure 3.6.: Scheme of how the distance between the projected hit position and the projected predicted position is used to group hits together.

which forms an ellipse in the parameter space spanned by  $\text{proj}_{\text{last hit}} - \text{proj}_{\text{first hit}}$  and  $\text{proj}_{\text{first hit}}$ . This is shown in Fig. 3.9. **majAxProj** and **minAxProj** are the major and minor axis of the ellipse respectively. The second group then starts with the hit which has the second smallest projection and forms a group with further hits which fulfill Eq. 3.13. See Fig. 3.6 for an illustration of this grouping. This procedure is repeated until no further groups can be formed. Each of these groups has to contain at least three hits in at least three different planes in  $\text{T}^*\text{T}$ .

**Fitting** A  $\chi^2$  fit is then performed for each group. The parametrisation in  $x$  and  $y$  is as follows:

$$x(z) = \text{offset}_x + \text{slope} \cdot (z_{\text{hit}} - z_{\text{TT proj. plane}}) \quad (3.14)$$

$$y(z) = \text{offset}_y, \quad (3.15)$$

with  $\text{offset}_x$ ,  $\text{slope}$  and  $\text{offset}_y$  as the fit parameters. This is a compromise as in principle the  $y$ -coordinate would also have to be fitted with a straight line. However, as the  $y$ -coordinate cannot be measured precisely in  $\text{T}^*\text{T}$ , this information cannot be included in the track fit. The three parameters are determined minimising the  $\chi^2/\text{ndof}$ <sup>11</sup> which can be derived as follows:

$$\begin{aligned} \chi^2/\text{ndof} = & \text{weight} \cdot (\text{offset}_x^2 + \text{offset}_y^2 + \\ & (\text{slope} \cdot (z_{\text{TT proj. plane}} - z_{\text{proj. point}}))^2) \\ & + \sum_{i \in \text{hits}} \frac{w_i \cdot \text{dist}_i^2}{n_{\text{hits}}} \end{aligned} \quad (3.16)$$

$$\begin{aligned} \text{with dist}_i = & \text{proj}_i - \text{offset}_x - \text{slope} \cdot (z_i - z_{\text{TT proj. plane}}) \\ & - \text{offset}_y \cdot \sin \theta. \end{aligned} \quad (3.17)$$

<sup>11</sup>ndof stands for «number of degrees of freedom» in the fit.



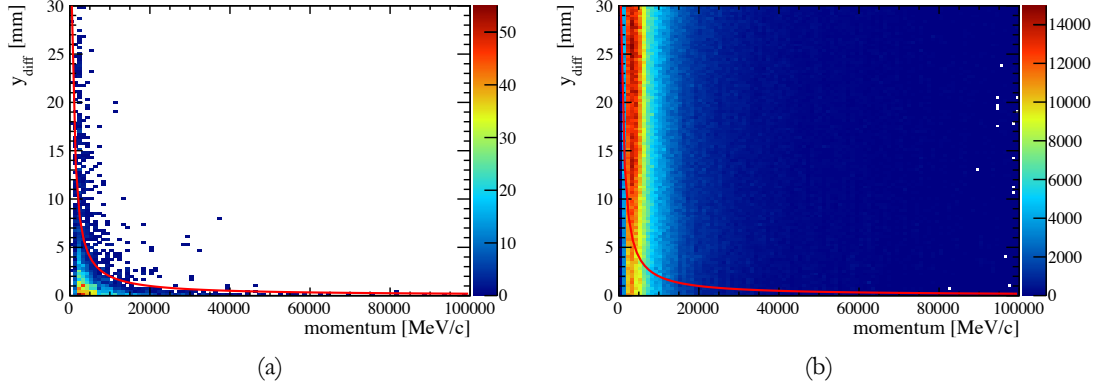


Figure 3.7.:  $y_{\text{diff}}$  as defined in Eq. 3.18 as a function of the momentum. Plotted in (a) for all true hits on a track, in (b) for all fake hits on a track. The red solid lines visualise the cut applied.

**weight** is a weighting factor set to a constant,  $w_i$  is the weight for an individual hit and  $\theta$  is the stereo angle of the TT layer. The first part of the expression probes the deviation from the extrapolated VELO- T station track while the second part is a measure of how well the hits agree with a straight line. Note that the fit is done in projected distances in  $x$ . Therefore, no absolute distances but only deviations from the predicted flight path are fitted.

If the  $\chi^2/\text{ndof}$  of the fit is larger than a given cut value, hits which give the largest contribution to the  $\chi^2/\text{ndof}$  (and are in a layer with at least two hits) are subsequently removed until only three hits are left or the  $\chi^2/\text{ndof}$  has dropped below the cut value. This procedure is repeated for every group of hits, where the final group is the one with the smallest  $\chi^2/\text{ndof}$  but the largest number of hits<sup>12</sup>.

### 3.6.2. Tuning of parameters

**Magnetic field parameters** The extrapolation between VELO and TT has two parameters that can be tuned: the  $z$ -position of the magnetic plane and  $\text{tt}_{\text{Par}}$ , a slope parameter. Calculating the absolute distance of the  $x$ -coordinates of the predicted hit positions between a full field extrapolation and the simplified extrapolation as written in Eq. 3.10 for every true<sup>13</sup> hit, a grid search (i.e. scanning both parameters simultaneously) was performed. The pair of parameters that yields the least distance was then chosen.

<sup>12</sup>A selected group of hits is replaced by a new one if the new group has a smaller  $\chi^2/\text{ndof}$  and at least as many hits as the previous one.

<sup>13</sup>«true» in this context means that it could be associated to a simulated particle.

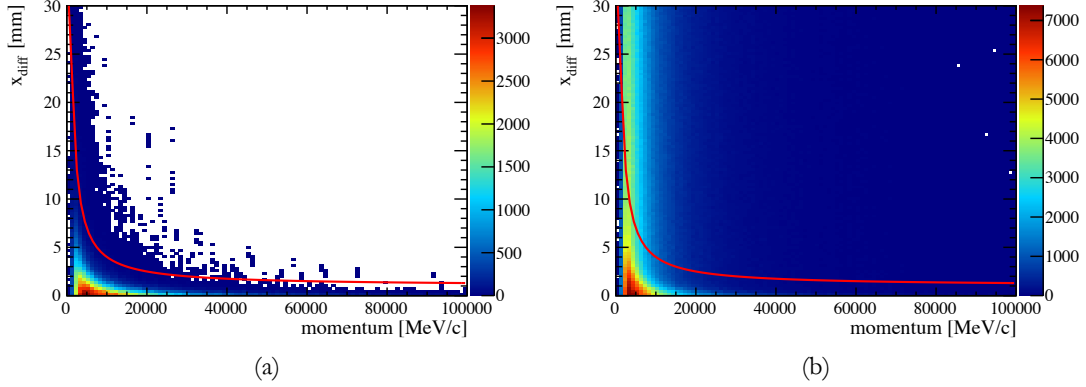


Figure 3.8.:  $x_{\text{diff}}$  as defined in Eq. 3.20 as a function of the momentum for all true hits on a track (a) and for all fake hits on a track (b). The red solid lines illustrates  $x_{\text{To1}}$ , the cut applied to the hits. The region above the line is rejected.

**y Tolerance** To optimise the tolerance in  $y$ , the distribution

$$y_{\text{diff}} = |y_{\text{pred}} - \frac{1}{2}(y_{\text{beg}} + y_{\text{end}})| - \frac{1}{2}(y_{\text{beg}} - y_{\text{end}}) \quad (3.18)$$

was drawn as a function of the momentum and the functional form of the cut value was chosen by eye, see Fig. 3.7.

**Projection point** If the projection point is chosen correctly, all the hits of a track should have the same projection (assuming all tracks scatter at the same point). It is therefore reasonable to choose the  $z$ -coordinate of this point such that the difference of the sum of all projections in T1a and the sum of all projections in T1b is as small as possible. A wrong projection point would lead to too large projections in T1a and too small projections in T1b or vice versa. Therefore the following quantity was minimised:

$$s = \sum_{i \in \text{true hits}} \left| \frac{\text{proj}_i^{1. \text{ layer}}}{n_{\text{hits in 1. layer}}} \right| + \left| \frac{\text{proj}_i^{2. \text{ layer}}}{n_{\text{hits in 2. layer}}} \right| - \left| \frac{\text{proj}_i^{3. \text{ layer}}}{n_{\text{hits in 3. layer}}} \right| - \left| \frac{\text{proj}_i^{4. \text{ layer}}}{n_{\text{hits in 4. layer}}} \right|, \quad (3.19)$$

and the corresponding value for the projection point was used.

**x Tolerance** The functional form of  $x_{\text{To1}}$  was chosen by plotting

$$x_{\text{diff}} = |x_{\text{pred}} - x_i| \quad (3.20)$$

for true and fake hits, see Fig. 3.8.

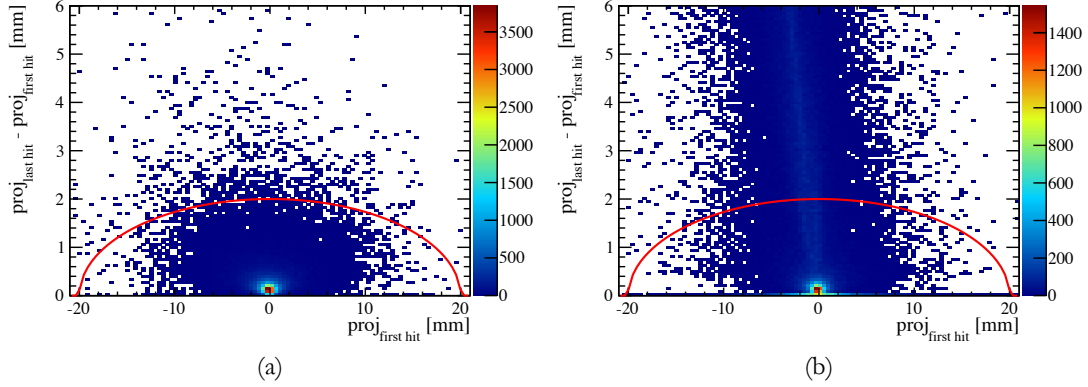


Figure 3.9.:  $\text{proj}_{\text{last hit}} - \text{proj}_{\text{first hit}}$  as a function of  $\text{proj}_{\text{first hit}}$ . Plotted in (a) for all true hits on track, in (b) for all fake hits on a track. The red solid lines visualise the cut applied.

**Projection Window** To parametrise the projection window, the first (i.e. smallest) and last (i.e. largest) projection of all true and fake hits on a track was calculated and plotted with respect to the first projection. An ellipsoidal form was chosen to be the most appropriate parametrisation for the upper cut value  $\text{proj}_{\text{max}}$ , see Fig. 3.9.

**Fit Tolerances** To determine the limit value for  $\chi^2/\text{ndof}$  called  $\chi^2_{\text{best}}$ , the  $\chi^2/\text{ndof}$  distribution was plotted as function of the momentum for all true hits for a given track, see Fig. 3.9.

All the parametrisations are summarised in Table 3.1.

**Note on the determined parameters** The procedure described in the last pages leads to a first estimate for the best parameters of the algorithm. However, in the end a fine-tuning by hand was performed. Although a single cut is well defined, it is difficult to precisely foresee the effect of a cut in, for example,  $x$  tolerance on the fit tolerance in the end. In addition, as most of the cuts are determined by eye, a certain arbitrariness is always included. Furthermore, the choice of tracks which are most «valuable» strongly depends on the physics goals. The solution chosen here was to achieve the best performance in terms of overall efficiency and purity (see Sect. 3.6.3).

### 3.6.3. Performance

The performance of the algorithm was measured using 10'000 simulated  $B^+ \rightarrow J/\psi K^+$  events. The composition of long tracks made with Forward Tracking and Track Matching was the same as used in the standard output of the track reconstruction. Two performance

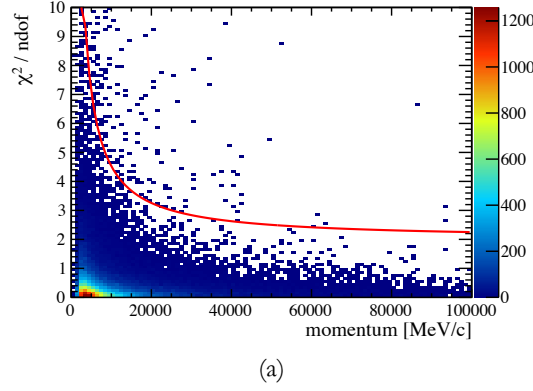


Figure 3.10.:  $\chi^2/\text{ndof}$  as defined in Eq. 3.14 as a function of the momentum, plotted for all true hits on a track. The red solid line visualises the cut applied.

Table 3.1.: Parametrisations in the PatAddTTCoord algorithm. Only parametrisations which are a function of another quantity are displayed.

Variable	Parametrisation
$y_{\text{Tol}}$	$y_{\text{Tol}} = \frac{y_{\text{TolSlope}}}{p}$
$x_{\text{Tol}}$	$x_{\text{Tol}} = x_{\text{TolOffset}} + \frac{x_{\text{TolSlope}}}{p}$
$\text{proj}_{\text{max}}$	$\text{proj}_{\text{max}} = \text{proj}_{\text{first hit}} + \sqrt{\text{majAxProj}^2 \cdot \left(1 - \frac{\text{proj}_{\text{first hit}}^2}{\text{minAxProj}^2}\right)}$
$\chi_{\text{best}}^2$	$\chi_{\text{best}}^2 = \text{maxChi2Tol} + \frac{\text{maxChi2Slope}}{p - \text{maxChi2POffset}}$

indicators are of interest: efficiency and purity. The efficiency indicates how good the algorithm is at adding the hits which belong to the track. The purity indicates how good the algorithm is in rejecting hits which do not belong to the track. The nomenclature is as follows: hits that could be matched to a hit caused by the corresponding simulated particle are called «true» hits; hits that were caused by a simulated particle are called «simulated» hits.<sup>14</sup> The efficiency and purity for a given long track are defined as follows:

$$\text{efficiency} = \frac{\# \text{ of true TT hits added to the long track}}{\# \text{ of simulated TT hits on the long track}} \quad (3.21)$$

$$\text{purity} = \frac{\# \text{ of true TT hits added to the long track}}{\# \text{ of TT hits added to the long track}} \quad (3.22)$$

<sup>14</sup>All true hits correspond to a simulated hit, but not all simulated hits have a true counterpart, due to reconstruction inefficiencies.

The average efficiency and purity numbers for all long tracks and for long tracks with at least three simulated hits are shown in Table 3.2. They are drawn as a function of the momentum in Fig. 3.11. The efficiency drops as a function of the momentum if no condition on the number of simulated TT hits is applied. This is mostly an acceptance effect as TT does not cover the highest regions in pseudorapidity. The effect vanishes if at least three simulated hits are required. The remaining inefficiency is largely due to inefficiencies in the pattern recognition.

Table 3.2.: Overall efficiency and purity for the PatAddTTCoord algorithm for all long tracks and for long tracks which have at least three simulated hits in TT. The numbers were obtained from a sample of 10'000 simulated  $B^+ \rightarrow J/\psi K^+$  events.

	all long tracks	long tracks, $\geq 3$ sim. TT hits
<b>Efficiency</b>	95.2 %	97.1 %
<b>Purity</b>	97.9 %	99.0 %

### 3.6.4. TT hit maps

To visualise how TT hits are added to long tracks, TT hit maps for all four layers were created with a sample of data taken in 2011 with magnetic field vector of the dipole magnet pointing downwards.<sup>15</sup> They are shown in Fig. 3.12 and reveal the internal structure of the TT with overlap regions and gaps. Note that these plots mainly show efficiencies that stem from geometrical properties of TT where the hit adding efficiency due to the performance of the pattern recognition is not the dominant contribution. To avoid having ghost tracks, only  $J/\psi \rightarrow \mu^+ \mu^-$  decays, where the invariant mass of both muons is inside a  $100 \text{ MeV}/c^2$  window of the nominal  $J/\psi$  mass, were considered.<sup>16</sup>

<sup>15</sup>TT was split into four layers only for these plots. No distinction was made between the individual  $z$ -position of the modules within one layer.

<sup>16</sup>Note that these plots only comprise tracks made with the Forward Tracking as the extrapolation from VELO to TT in the Track Matching contained a bug for the 2011 data taking period.

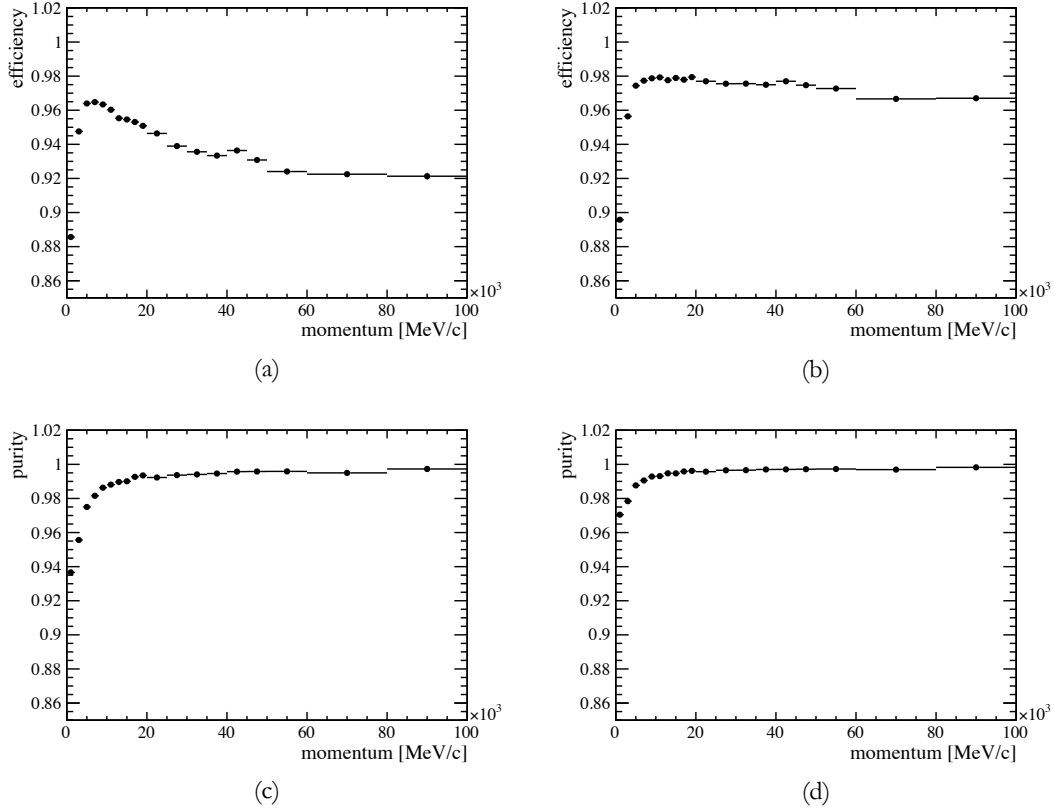


Figure 3.11.: Efficiency and purity of adding a TT hit to a long track as a function of the momentum, derived from a simulation of  $10'000 B^+ \rightarrow J/\psi K^+$  events. (a) and (c) are the efficiency and purity with no restriction on the number of simulated TT hits. (b) and (d) require at least three simulated TT hits.

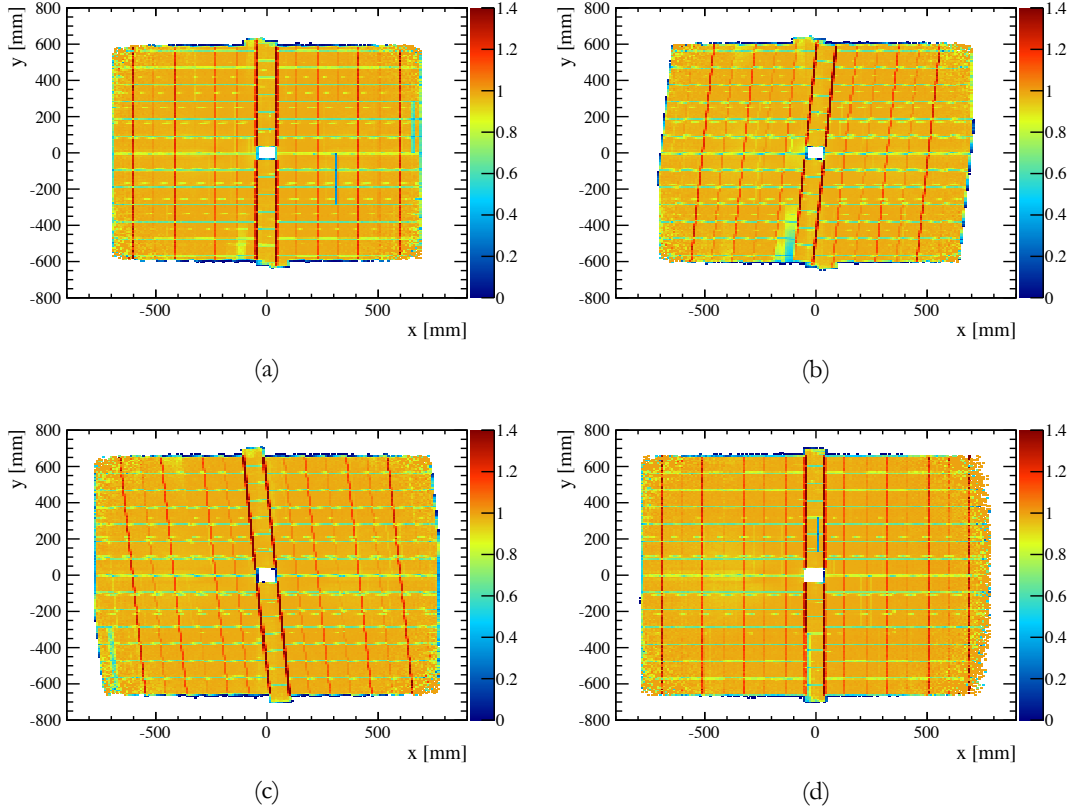


Figure 3.12.: Average number of TT hits added to long tracks in (a) TTaX, (b) TTaU, (c) TTbV, (d) TTbX in a  $J/\psi \rightarrow \mu^+\mu^-$  sample from the 2011 data taking (with downwards magnetic field). Clearly visible are the overlap regions with a higher average and some dead channels with a lower average.





## 4. Tracking efficiency in LHCb

In the previous chapter it was shown how hits are bundled together and thus tracks are formed. It is in the nature of the procedure that this «track finding» is not fully efficient, meaning that not all flight paths of particles traversing the LHCb detector can be reconstructed. The knowledge of this track reconstruction efficiency or «tracking efficiency» however is important for every physics analysis that needs to compute the branching fraction, cross-section or reconstruction efficiency of a decay. In this chapter, the tracking efficiency for long tracks is explored in depth. A novel method is presented how it can be determined from collision data using the muon system and the TT, and how it can be applied for different momentum regimes. The method is then compared with an alternative one, using an approach which combines the tracking efficiency measured in the VELO and the T stations.

The task of determining the tracking efficiency is no standard procedure and depends very much on the type of the detector in question. The description given in this chapter is therefore rather technical and tries to cover all possible effects and idiosyncrasies of the technique which was developed for LHCb.

### 4.1. Approaches to measure the tracking efficiency

Several ways exist to compute the tracking efficiency: the simplest is to consider simulated events, where the ratio of reconstructed particles over all particles traversing the detector can be computed. However it is a priori not known how well simulation and collision data agree. A further method is to measure ratios of well-known branching fractions, such as  $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$  and  $D^0 \rightarrow K^- \pi^+$ . The different yields of these two decays then only stem from the different branching fractions and the reconstruction efficiency for the two additional pions. The reconstruction efficiency comprises on the one hand the efficiency for a particle to reach all the relevant tracking stations and on the other hand the efficiency of the flight path to be reconstructed as a track. For an application of this method see Ref. [35].

Another method is to measure the tracking efficiency in a direct way. This can be accomplished by checking for the presence of a track using a subdetector which was not used in the standard track finding procedure, *e.g.* the muon system. Alternatively, the efficiency for a given subdetector can be measured by excluding it from the track finding procedure and using other subdetectors to probe it. In both cases one has to make sure no ghost tracks are

considered. This can, for example, be achieved by selecting resonances with a large branching fraction to two-prong decays, where one track is fully reconstructed («tag-track»), while the other only partially («probe-track»). One then has to make sure that the partially reconstructed track comprises enough information to allow selecting the resonance cleanly enough by computing the invariant mass of the combination of tag and probe-track. This procedure is commonly referred to as «tag-and-probe». An illustration for tag-and-probe for a combination of a long track and a «MuonTT» track (see next section) is given in Fig. 4.1.

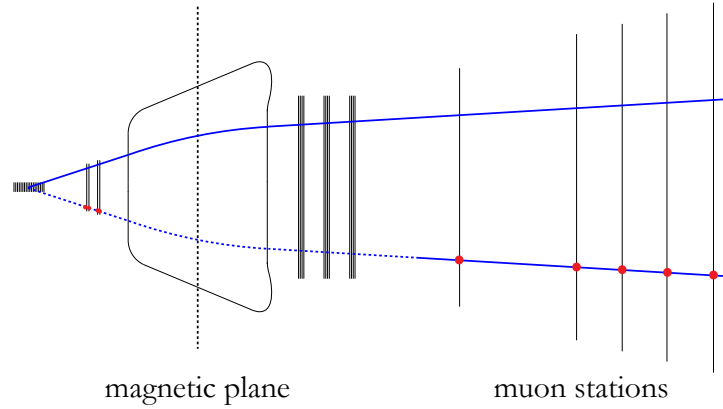


Figure 4.1.: Illustration of the tag-and-probe approach to measure the tracking efficiency. The upper blue line is a fully reconstructed long track and the lower dashed line a partially reconstructed track (MuonTT track) that serves as a probe to check for the existence of a long track. Both tracks together form the decay products of a resonance, *e.g.* a  $J/\psi$  or  $Z^0$ .

Two methods were used to determine the tracking efficiency in the 2010 and 2011 collision data in LHCb [100]. They are referred to as «Combined Method» — consisting of «VELO Method» and «T Station Method» — and «Long Method». The Long Method is explained in detail in the following pages. It uses  $J/\psi \rightarrow \mu^+ \mu^-$  decays to access the low momentum region and  $Z^0 \rightarrow \mu^+ \mu^-$  decays to access the high momentum region. The Combined Method was developed independently and will be described in Sect. 4.7.

## 4.2. Construction of MuonTT tracks

### 4.2.1. Muon standalone tracks

The muon system in LHCb was designed to provide particle identification for muons. It is built similarly to a tracking detector and can in principle be used as such. It provides a 3-dimensional position for hits left by muons, *e.g.* allowing the reconstruction of the

flight path of cosmic muons for efficiency studies [101]. However, its coarse granularity of the multi-wire proportional chambers and triple-GEM detectors and the large amount of multiple scattering induced by the iron-absorber layers significantly reduce the momentum resolution. Despite these shortcomings the muon system serves as the starting point for the formation of «MuonTT tracks» which will be used in the Long Method.

#### 4.2.2. Formation of muon standalone tracks

Different algorithms exist to form muon standalone tracks: due to the simple geometry of the muon system and the lack of a magnetic field they all have a similar structure. The algorithm used for the work presented in this thesis was developed for the reconstruction of cosmic muons and the monitoring of the performance of the muon stations. Two flavours of this algorithm exist: One performing a pattern recognition using a neural network [102] and one using a cut based pattern recognition. The first is too slow to be used for collision data while the second performs equally well while being significantly faster; this version will be explained below [103].

**Seeding & Pattern recognition** A muon station is chosen as the seed station. The choice has to be a balance between purity (increases towards M5) and efficiency (increases towards M1). All the hits in the seed station are considered starting points for seeds. The hit position is «extrapolated» using a linear function:

$$x_{SS-1} = x_{SS} \cdot \frac{z_{SS-1}}{z_{SS}}, \quad (4.1)$$

and, where,  $x_{SS,SS-1}$  and  $z_{SS,SS-1}$  are the  $x$ - and  $z$ -coordinate of the hit in the seed station and the station before the seed station, respectively. The extrapolation in  $y$  proceeds similarly. The seeding always proceeds in the negative  $z$  direction. This linear extrapolation corresponds to a straight track coming from the interaction point, defined at the origin of the coordinate system. This provides an accurate description in the  $y$ - $z$  plane whilst it is strictly only true for tracks of infinite momentum in the  $x$ - $z$ -plane. Hits are then searched for in the vicinity of this extrapolated position inside a suitable window, where the window size depends on the muon station in question. If a hit is found, both hits together form a seed track. If more than one hit is found inside the search window, the hit closest to the extrapolated position is taken to form a seed track. The seed track is then extrapolated into the next station closer to the magnet, where the extrapolation in  $x$  is along direction of the seed track and in  $y$  towards the interaction point. In this station the hit which is closest to the extrapolated position is again added. This procedure continues until M1 is reached. The search then starts again with a new seed. Note that hits in muon stations downstream of the seed station are not considered and that no track is formed if no (compatible) hit can be found in a station closer to the magnet than the seed station.

**Fitting:** The muon tracks are fitted using a  $\chi^2$  minimisation in the  $x$ - $z$ - and the  $y$ - $z$ -projection to all hits selected by the pattern recognition. The formula for this linear track fit and its derivation are given in Appendix A.

Although there is no magnetic field in the muon system, a straight line parametrisation is not necessarily correct, as the iron absorbers induce large amounts of multiple scattering not taken into account in the straight line fit. This shortcoming is not considered any further in the context of the tracking efficiency measurement presented in this thesis. Possible solutions are described in Ref. [104].

### 4.2.3. Momentum estimate

As the muon stations are far outside the magnetic field, no momentum measurement can be made using only the muon tracks. However, assuming that the muon track originated at the interaction point a momentum estimate can be given using the same optical model as in the Forward Tracking, see Sect. 3.5.1: the track in the muon system defines the slope and position, while the primary vertex acts as the additional point for the momentum determination. The momentum is calculated using the following formula which is a consequence of the Lorentz force [105]:

$$\Delta \vec{p} = q \cdot \int \vec{B} \times d\vec{l} \quad (4.2)$$

As the bending primarily occurs in the  $x$ - $z$ -plane, Eq. 4.2 can be simplified to:

$$\Delta p_x = p_{x,f} - p_{x,i} = q \cdot \int (\vec{B} \times d\vec{l})_x, \quad (4.3)$$

where  $p_{x,i}$  and  $p_{x,f}$  stand for the momentum component in the  $x$ -direction before and after traversing the magnetic field, respectively. These two quantities can be related to the absolute momentum:

$$\Delta p_x = p_{x,f} - p_{x,i} = p \cdot \left( \frac{t_{x,f}}{\sqrt{1 + t_{x,f}^2 + t_{y,f}^2}} - \frac{t_{x,i}}{\sqrt{1 + t_{x,i}^2 + t_{y,i}^2}} \right) \quad (4.4)$$

with  $t_{x,j} = \left(\frac{dx}{dz}\right)_j$  and  $t_{y,j} = \left(\frac{dy}{dz}\right)_j$  denoting the slopes of the track. An illustration for this «p-kick» method is given in Fig. 4.2. Thus the problem is reduced to a calculation of the integral in Eq. 4.3 and the slope of the track in the muon system and at the primary vertex. While the first slope is directly taken from the muon track, the latter is calculated by linearly extrapolating the muon track to the magnetic plane and then determining the slopes from the difference in the  $x$ -,  $y$ - and  $z$ -coordinates to the primary vertex.

The integral in Eq. 4.3 is calculated using the magnetic field map and stepping along an imaginary track with a given fixed slope. The integral does not depend on the actual path of the track for the purpose of this measurement and it is only calculated once and not

for every muon track. The sign of the charge of the particle corresponding to the track is determined by considering the slopes at the vertex and the muon system and the polarity of the magnetic field.

It should be noted that the momentum determined with the outlined method shows a certain bias depending on  $t_x^2$ , due to inhomogeneities in the magnetic field. A (parabolic) correction is therefore applied to the obtained momentum estimate.

To obtain a better precision on the momentum, an additional procedure is performed for the tracking efficiency studies: The track state, constructed out of the muon standalone track and the momentum estimate, is extrapolated to the primary vertex, using a full extrapolation<sup>1</sup>. The difference in  $x$  from zero is then used to adapt  $q/p$  for the state in the muon system. This procedure is repeated until the deviation in the  $x$ -coordinate from zero at the primary vertex is smaller than a given value. The procedure normally converges after a few iterations.

The momentum resolution for muon standalone tracks and the dimuon spectrum, formed out of a long track and a muon standalone track on a simulated  $J/\psi \rightarrow \mu^+ \mu^-$  sample, are shown in Fig. 4.3. The data points were fitted with a double Gaussian<sup>2</sup> probability density function (PDF) and an exponential distribution

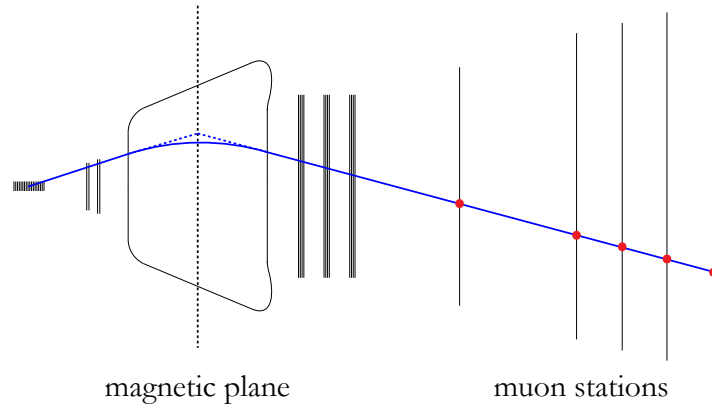


Figure 4.2.: Illustration of the p-kick method to determine the momentum of a muon standalone track. The blue line is the real flight path, the dashed blue line is the linear extrapolation to the magnetic plane and the red dots are the hits in the muon stations.

<sup>1</sup>A full extrapolation uses the detailed magnetic field map and does not rely on a simplified approach like a p-kick.

<sup>2</sup>A double Gaussian is the sum of two Gaussian PDFs with a common mean.

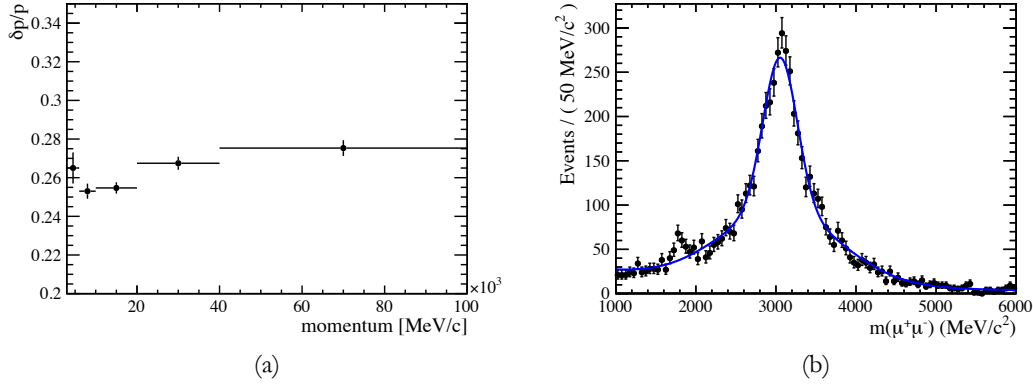


Figure 4.3.: Performance of the muon standalone track reconstruction, evaluated on 15'000 simulated  $J/\psi \rightarrow \mu^+\mu^-$  events. In (a) the momentum resolution of the standalone track is shown while in (b) the combination of a fully reconstructed long track and a muon standalone track of opposite charge is shown. A clear mass peak around the nominal  $J/\psi$  mass can be observed. The data points were fitted with a double Gaussian PDF and an exponential distribution. The resolution of the narrower component is approximately 210 MeV/c<sup>2</sup>.

#### 4.2.4. MuonTT tracks

The muon standalone tracks can be combined with hits in the TT, following<sup>3</sup> what is explained in Sect. 3.6. However, as the momentum and the position are much less precisely defined than from a VELO-T-station track, the search windows have to be widened in order to find the correct hits. The hits in the muon system and TT are then fitted with the Kalman filter to form a track. The momentum resolution for MuonTT tracks and the dimuon spectrum, formed out of a long and a MuonTT track on a simulated  $J/\psi \rightarrow \mu^+\mu^-$  sample, are shown in Fig. 4.4. The data points were fitted with two Crystal Ball PDFs [106] with a common mean and an exponential distribution. The mass and momentum resolution improves by a factor of 7 - 10 when adding TT hits to the muon standalone track compared to the muon standalone track alone, see Table 4.1.

<sup>3</sup>Technically the state in the muon system is extrapolated to the VELO. This state is then extrapolated back to TT to add hits following the standard algorithm.

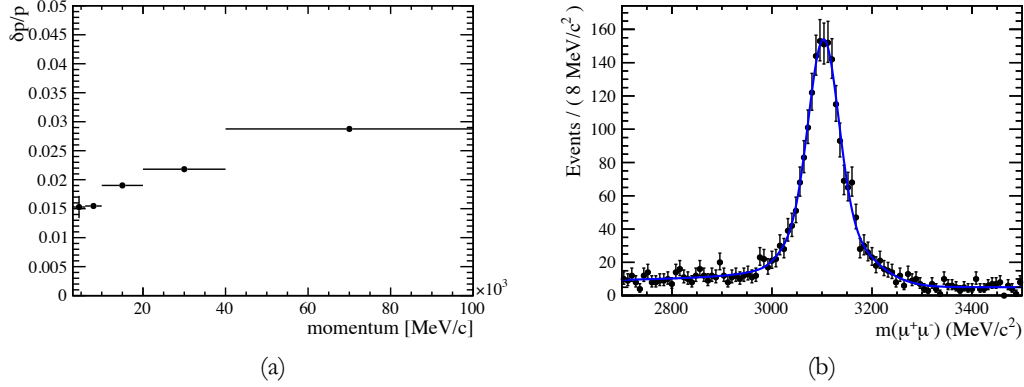


Figure 4.4.: Performance of the MuonTT tracking, evaluated on 15'000 simulated  $J/\psi \rightarrow \mu^+\mu^-$  events. In (a) the momentum resolution of the MuonTT track is shown while in (b) the combination of a fully reconstructed long track and a MuonTT track of opposite charge is shown. A clear mass peak around the nominal  $J/\psi$  mass can be observed. The data points were fitted with two Crystal Ball PDFs [106] with a common mean and an exponential distribution. The resolution of the narrower component is approximately  $30 \text{ MeV}/c^2$ .

Table 4.1.: Comparison of the mean and the narrow component of the mass resolution for  $J/\psi \rightarrow \mu^+\mu^-$  for a combination of a long track with a muon standalone track, with a MuonTT track and with a long track. The values are only approximate hence no uncertainty is given. The value for the long – long combination was determined on a  $B^0 \rightarrow J/\psi K^{*0}$  sample from 2011 collision data with two Crystal Ball PDFs with a common mean for the dimuon mass distribution. Note that the mean value of the long - muon standalone combination has a bias towards low values for the mean of the  $J/\psi$  mass.

mode	mean [ $\text{MeV}/c^2$ ]	$\sigma_{\text{narrow}}$ [ $\text{MeV}/c^2$ ]
long - muon standalone	3053	213
long - MuonTT	3103	30
long - long	3100	12

### 4.3. Reconstruction of dimuon resonances in collider data

In order to apply the tag-and-probe approach to collision data, a handful of precautions have to be taken. Most importantly biases from the trigger are a concern: As the HLT already does a track reconstruction, one has to make sure the trigger decision does not influence the tracking efficiency measurement. This is achieved by requiring the tag-track to be aligned to the track that fired the trigger in HLT1 and HLT2 [73].<sup>4</sup> Only trigger decisions relying on single tracks can therefore be used for this tracking efficiency measurement. The level-0 trigger only reconstructs muon standalone tracks and does therefore introduce a bias. The trigger decisions that were used for low- and high transverse momentum tracks are given in Table 4.2 and 4.3.

Table 4.2.: Trigger lines and selections in HLT1 and HLT2 used for tracking efficiency measurements for low transverse momentum tracks.

Name	Prescale (2010/2011)	Selection (2010)	Selection (2011)
Hlt1TrackMuon	1.0/1.0	IP > 0.11 mm	IP > 0.1 mm
		–	$\chi_{\text{IP}}^2 > 16^5$
		$p > 8 \text{ GeV}/c$	$p > 8 \text{ GeV}/c$
		$p_{\text{T}} > 0.8 \text{ GeV}/c$	$p_{\text{T}} > 1.0 \text{ GeV}/c$
		$\chi_{\text{track}}^2/\text{ndof} < 10$	$\chi_{\text{track}}^2/\text{ndof} < 2$ $\chi_{\text{muon}}^2 < 25^6$
Hlt2SingleMuon	0.02/0.5	IP > 0.08 mm	IP > 0.5 mm
		–	$\chi_{\text{IP}}^2 > 200$
		$p_{\text{T}} > 1.3 \text{ GeV}/c$	$p_{\text{T}} > 1.3 \text{ GeV}/c$
		$\chi_{\text{track}}^2/\text{ndof} < 2$	$\chi_{\text{track}}^2/\text{ndof} < 2$
		isMuon = true <sup>7</sup>	isMuon = true

<sup>4</sup>Although the track reconstruction in the HLT and offline are similar, they are not identical. Therefore a quality criterion based on hits in the subdetectors is in place to «match» offline tracks and tracks in the HLT.

<sup>5</sup>The impact parameter  $\chi^2$  is a measure of how compatible the track is with originating at the primary vertex.

<sup>6</sup> $\chi_{\text{muon}}^2$  is the matching quality of a VELO and a muon standalone track in HLT1. It is described in detail in Ref. [73].

<sup>7</sup>They need to fulfill the isMuon criterion: For a momentum with  $3 < p < 6 \text{ GeV}/c$ , hits need to be found in M2 and M3, for  $6 < p < 10 \text{ GeV}/c$ , hits need to be found in M2, M3 and M4 or M5, and for  $10 < p$ , hits need to be present in muon stations M2, M3, M4 and M5.



Table 4.3.: Trigger lines and selections in HLT1 and HLT2 used for tracking efficiency measurements for high transverse momentum tracks.

Name	Prescale	Selection (2010)
Hlt1SingleMuonNoIPL0	0.2	$p_T > 1.35/1.8 \text{ GeV}/c$ $\chi_{\text{track}}^2/\text{ndof} < 50/10$ <sup>8</sup> $\chi_{\text{muon}}^2 < 100/16$ <sup>9</sup>
-----		
Hlt1SingleMuonNoIPL0HighPT	1.0	$p_T > 5 \text{ GeV}/c$ $\chi_{\text{track}}^2/\text{ndof} < 50/10$ $\chi_{\text{muon}}^2 < 100/16$
Hlt2SingleHighPTMuon	1.0	$p_T > 10 \text{ GeV}/c$

A second (potential) bias comes from the «stripping»<sup>10</sup>. Although these selections are done offline, they behave similarly to the trigger, as the end-user cannot access data not coming through the stripping. The decisions in the trigger are based on simple objects (*e.g.* tracks) and preexisting trigger lines could be exploited. The stripping is more complex. For high transverse momentum tracks, a single muon stripping line exists (for  $W$  analyses). For low  $p_T$  tracks, nothing comparable to a trigger line can be used. The construction of MuonTT tracks is therefore done in the stripping.

#### 4.3.1. Constructing a dimuon resonance

The workflow to reconstruct a dimuon resonance is similar for both, high- and low- $p_T$ . The only difference is the place where it is executed (offline or stripping).

The basic idea is to reconstruct every resonance twice: once with a positive muon and once with a negative muon as the tag-track. The workflow for the first case is described below:

- 1 Retrieve all particles in an event with a positive charge and identified as muons.<sup>11</sup>
- 2 Require them to be aligned with the track that has fired the single muon trigger in HLT1 and HLT2.

<sup>8</sup>The value was tightened for later runs in the 2010 data taking.

<sup>9</sup>The value was tightened for later runs in the 2010 data taking. This matching criterion for muon hits is not identical to the one used in 2011, it is described in Ref. [107].

<sup>10</sup>«Stripping» is the term for offline selections which are run centrally to reduce the amount of data the user has to analyse for a specific decay channel.

<sup>11</sup>They need to fulfill the isMuon criterion. Furthermore a cut on a combined likelihood for muon identification using different subdetectors is applied. For details on these quantities see Sect. 6.2.2 and Ref. [108].

- 3 Do the pattern recognition for MuonTT tracks, assign a muon mass hypothesis to them. Select only the ones with a negative charge.
- 4 Combine the two tracks in a vertex fit and apply a criterion on the quality of the fit.

After these steps, a clear mass peak is visible in the dimuon mass spectrum, which is then used for the measurement. The task is to select the  $J/\psi$  cleanly enough, to define a criterion for which a track is called «found» and to analyse different systematic effects. All this will be discussed below.

## 4.4. Prerequisites for a tracking efficiency measurement

As the tracking efficiency measurement using the Long Method was uncharted territory, a handful of cross-checks and optimisations were performed to make sure its performance is satisfactory.

### 4.4.1. Association

To measure the tracking efficiency, the presence of a long track corresponding to the MuonTT track needs to be checked. The tracking efficiency can then be formulated as:

$$\varepsilon_{\text{track}} = \frac{\text{\# of MuonTT tracks associated to a long track}}{\text{\# of MuonTT tracks}}. \quad (4.5)$$

The association is done by checking the overlap of hits of both tracks in the muon system<sup>12</sup> and possibly TT. The fact that not all long tracks have TT hits leads to a complication. In principle one would need to disentangle the lack of hits in TT due to inefficiencies in the pattern recognition, *e.g.* only adding hits if three or more are found; and due to acceptance effects, *e.g.* a track could have only passed TT at the very edge of a layer or gone through a gap between the sensors. As this distinction would be very difficult or maybe not even feasible, a pragmatic approach was chosen: the lack of TT hits on a long track was considered to be entirely due to the pattern recognition. This allows the MuonTT track to be taken as a probe for every long track in the event.

In practice, the overlap was checked as follows: both tracks, MuonTT and long, need to share a certain fraction of hits in the muon stations. In addition, if the long track which is compared to the MuonTT track has TT hits, both tracks also need to share a certain fraction of hits in TT.

---

<sup>12</sup>Technically a long track does not have hits in the muon stations, as these hits are not used in the pattern recognition to form a long track. They are the hits used to identify the particle as a muon. The difference is dropped in the text for simplicity.

These fractions have to be determined on simulation. The following procedure is performed: In a sample of simulated  $X \rightarrow \mu^+ \mu^-$  decays, where  $X$  is a  $Z^0$  or  $J/\psi$ , a track corresponding to a muon from the resonance is selected and its overlap compared with the MuonTT track of the same charge. As by definition these tracks have to match, their overlap fraction can be used to define a cut for the application in collision data. Furthermore, it is also checked how good the overlap between the MuonTT track and a random track, not belonging to the muon from the resonance, is where the worst possible track (*i.e.* having the largest overlap) is considered.

The comparison of these two «matchings» is shown in Figs. 4.5 and 4.6. From the two studies, the required overlap fraction is set to the values shown in Table 4.4. There is not always a clear separation between the fractions for the correctly or incorrectly matched track when considering the muon system alone. However this deficiency is largely reduced by the requirement to pass the corresponding association fractions in both TT and the muon system. The ambiguity in the association in the muon system then only affects long tracks without TT hits. This forms a small portion of the samples used for data analysis.

Table 4.4.: Overlap fractions in the muon system and TT for tracking efficiency measurements using  $Z^0 \rightarrow \mu^+ \mu^-$  and  $J/\psi \rightarrow \mu^+ \mu^-$ .

Decay	fraction in muon system	fraction in TT
$J/\psi \rightarrow \mu^+ \mu^-$	70%	60%
$Z^0 \rightarrow \mu^+ \mu^-$	40% <sup>13</sup>	60%

#### 4.4.2. General crosschecks

**Correlations** For an unbiased tracking efficiency measurement, no correlation must exist between forming a MuonTT track and a long track out of a muon in a given event. A bias could be present if large multiple scattering would occur in the detector. In this case, the pattern recognition in the T stations would fail to find the right hits for the long track while the pattern recognition in TT would fail for the MuonTT track. The method presented for tracking efficiency determination would not account for such kind of an effect. The possible size of this correlation was checked on simulated  $Z^0 \rightarrow \mu^+ \mu^-$  data and is very small; it is not considered any further. The result is shown in Fig. 4.7.

A similar effect may be induced by hadronic interactions for kaons and pions. This is discussed further in Sect. 4.8.3. Furthermore it is checked that the presence of one reconstructed muon does not influence the efficiency to reconstruct the second one.

<sup>13</sup>In retrospect, the cut for the overlap in the muon system for  $Z^0 \rightarrow \mu^+ \mu^-$  seems rather loose. However, as for almost all the tracks also the overlap in TT has to be larger than 60%, the pollution from incorrectly matched tracks is negligible.

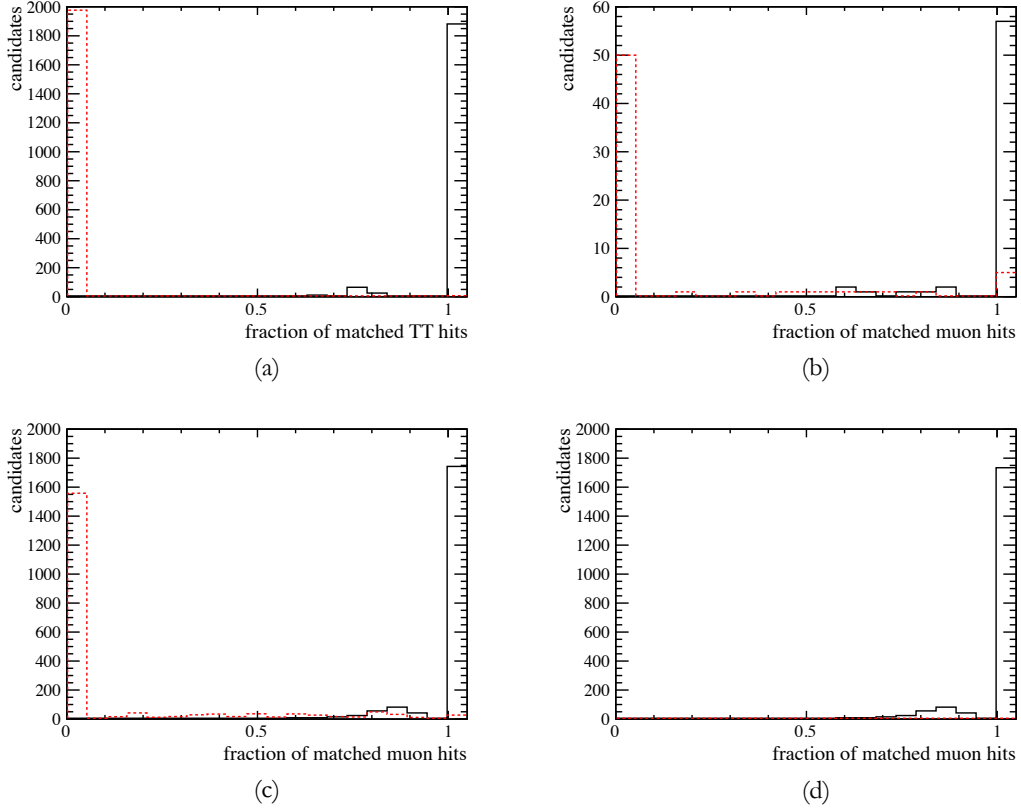


Figure 4.5.: Overlap between MuonTT track and long track for simulated  $Z^0 \rightarrow \mu^+\mu^-$  events in TT (a), in the muon system if the long track does not have TT hits (b), in the muon system if the long track has TT hits (c) and in the muon system if more than 60% of the hits in TT need to agree (d). In dashed red is the overlap of tracks which have the largest overlap in muon and TT hits, but do not stem from a  $Z^0$ , while in black is the overlap of tracks that stem from a  $Z^0$ . The conditions were chosen such that every event contains a fully reconstructed  $Z^0 \rightarrow \mu^+\mu^-$  event which could be matched to simulation.

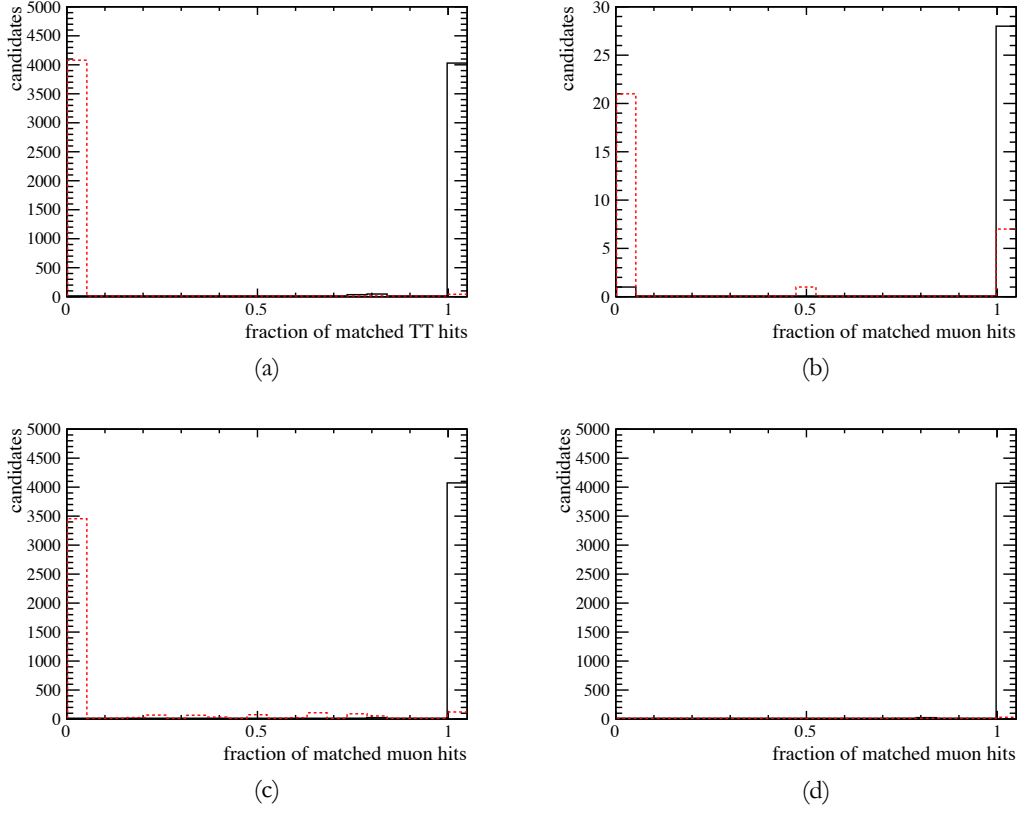


Figure 4.6.: Overlap between MuonTT track and long track for simulated  $J/\psi \rightarrow \mu^+\mu^-$  events in TT (a), in the muon system if the long track does not have TT hits (b), in the muon system if the long track has TT hits (c) and in the muon system if more than 60% of the hits in TT need to agree (d). In dashed red is the overlap of tracks which have the largest overlap in muon and TT hits, but do not stem from a  $J/\psi$ , while in black is the overlap of tracks that stem from a  $J/\psi$ . The conditions were chosen such that every event contains a fully reconstructed  $J/\psi \rightarrow \mu^+\mu^-$  event which could be matched to simulation.

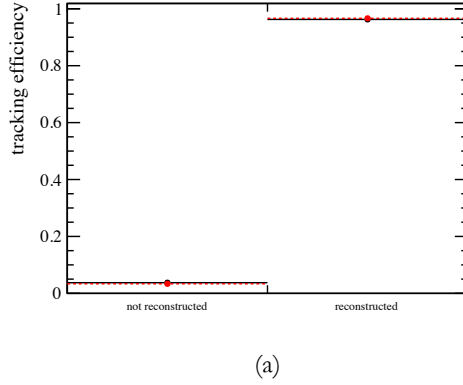


Figure 4.7.: Correlation check for MuonTT tracks on simulation. In dashed red is the fraction of tracks reconstructed from a  $Z^0 \rightarrow \mu^+\mu^-$  decay when a MuonTT track was reconstructed in the event, in black is the fraction of tracks reconstructed from a  $Z^0 \rightarrow \mu^+\mu^-$  decay when no such criterion is applied.

**Phase-space distance of associated tracks** In order to cross-check the association as described in Sect. 4.4.1, the «distance» between a MuonTT track and long tracks was visualised in the  $(\Delta R - \frac{\Delta p}{p})$  plane, where  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ . Correctly associated tracks should cluster around (0,0) while other tracks should be more randomly distributed<sup>14</sup>. This behaviour was confirmed on a simulated  $Z^0 \rightarrow \mu^+\mu^-$  sample and visualised in Fig. 4.8.

In principle, the distance in the  $(\Delta R - \frac{\Delta p}{p})$  plane can also be used to directly associate the probe-tracks to the reconstructed (long) tracks and give a measurement of the tracking efficiency (*e.g.* as done in Ref. [109]). For LHCb, the matching via hits in subdetectors was deemed more appropriate.

## 4.5. Tracking efficiency for high transverse momentum tracks

Understanding the tracking efficiency for high transverse momentum tracks is a crucial ingredient for cross-section measurements of  $W$  and  $Z^0$  bosons [110]. The idea is to use the tag-and-probe approach on reconstructed  $Z^0 \rightarrow \mu^+\mu^-$  decays where one muon corresponds to a long track and the other one to a MuonTT track. The following determinations are all done on the  $37.5 \text{ pb}^{-1}$  of collision data recorded in 2010.

<sup>14</sup>Considering  $\Delta R$  alone may not be enough:  $\Delta R$  could be the same for two different tracks with the same flight direction after their origin.

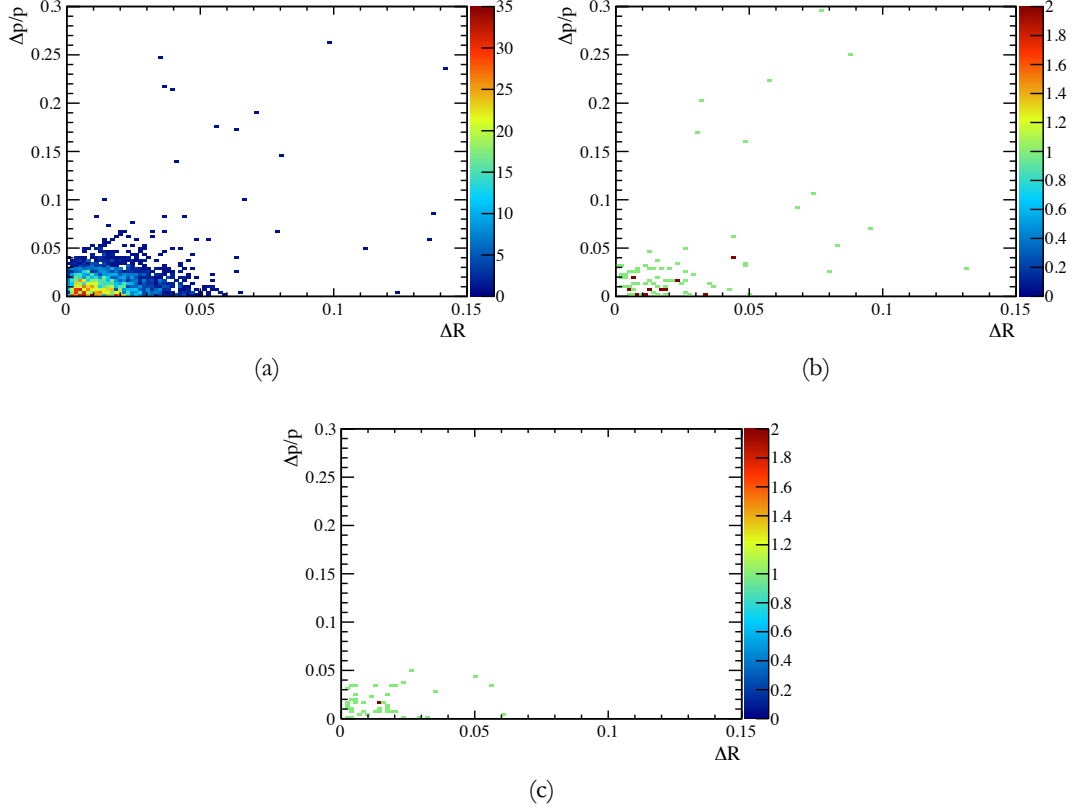


Figure 4.8.: Distributions of tracks in the  $(\Delta R - \frac{\Delta p}{p})$  plane of MuonTT track and associated long track. Figure (a) shows the distribution of long tracks which were associated to a MuonTT track. Figure (b) shows the distribution of long tracks which have the smallest  $\Delta R$  in an event. The long track that was associated to a MuonTT track was excluded. Figure (c) shows the distribution of long tracks which have the smallest  $\Delta R$  in an event. No long track was associated to a MuonTT track in these events.

### 4.5.1. Event selection

In order to obtain a reasonably clean, but unbiased sample of  $Z^0$  for the efficiency measurement, some selection cuts are applied, see Table 4.5. Note that no cuts on the probe-muon are applied, except for cuts defining the region of interest for the measurement, and the cuts on the tag-muon are chosen not to bias the probe-muon<sup>15</sup>. The resulting invariant mass distribution of the two muons is shown in Fig. 4.10.

Although a small amount of background cannot be excluded after these cuts, it is neglected for several reasons. First, considering the mass distributions in Fig. 4.10, no conclusive statement or estimation of the amount of background is possible. Second, only background coming from ghosts would be problematic, as real tracks in the background should have the same efficiency as in the signal. Finally, the precision with which the efficiency can be determined given the amount of data makes the contribution from ghost tracks in the background negligible.

Table 4.5.: Selection cuts for the  $Z^0$  resonance for the tracking efficiency measurement on 2010 collision data. The cuts on the probe-muon only define the region of interest in which the efficiency is measured.

Variable	Value
$p_T$ tag-muon	$> 30 \text{ GeV}/c$
$\chi^2_{\text{track}}/\text{ndof}$ tag-muon	$< 3$
isMuon tag-muon	true
$\text{DLL}_{\mu\pi}$ tag-muon <sup>16</sup>	$> -2$
$p_T$ probe-muon	$> 20 \text{ GeV}/c$
$\eta$ probe-muon	$2 < \eta < 4.5$
$Z^0 \chi^2_{\text{vtx}}$	$< 4$
$Z^0$ mass window	$20 \text{ GeV}/c^2$
# $Z^0$ candidates/event	1

### 4.5.2. Checks on simulated data

The tracking efficiency measured with the Long Method was compared to the tracking efficiency on the simulation level to check the presence of possible systematic effects. In simulation, it is possible to obtain the tracking efficiency directly by checking if a particle trajectory which is «reconstructible» is also reconstructed as a track by the tracking algorithms.

<sup>15</sup>As could, for example, be the case for cuts on the impact parameter.

<sup>16</sup> $\text{DLL}_{\mu\pi}$  is explained in Sect. 6.2.2.



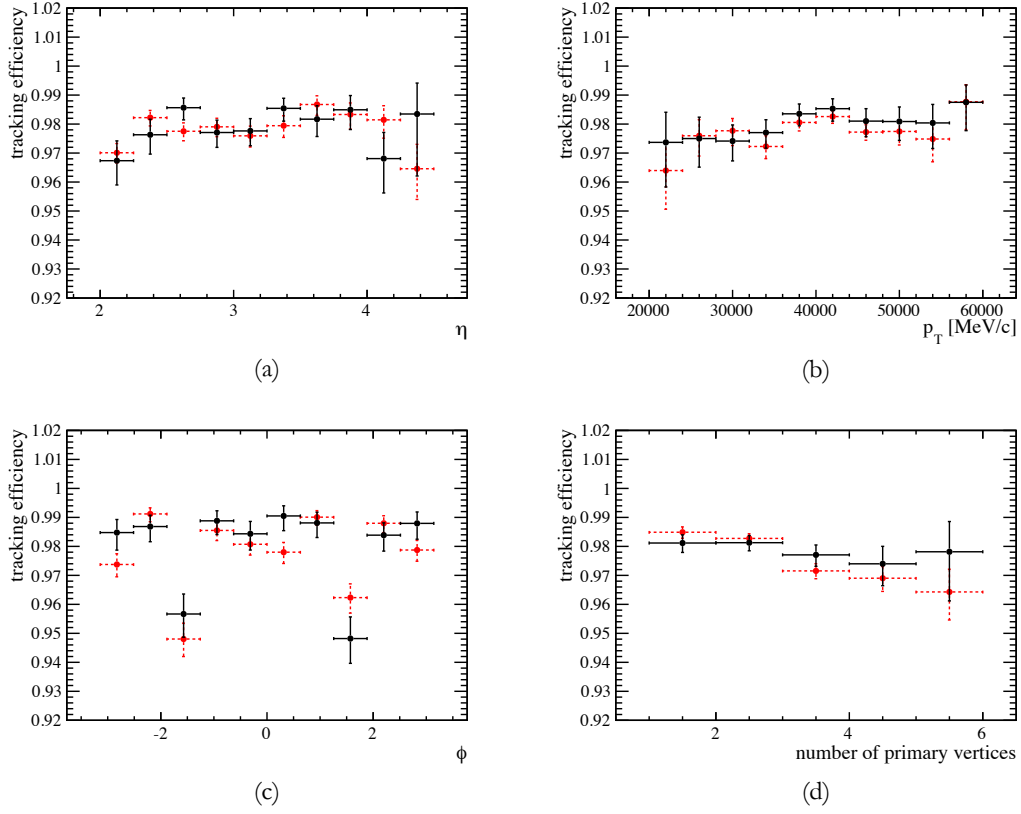


Figure 4.9.: Comparison between the tracking efficiency determined with the hit-based method (red dashed data points) and with the tag-and-probe approach (black solid data points) on a sample of simulated  $Z^0 \rightarrow \mu^+ \mu^-$  events, binned in pseudorapidity (a), transverse momentum (b), polar angle (c) and number of reconstructed primary vertices (d). Both methods show a good agreement within  $\approx 1\%$ . The two dips in the  $\phi$  distribution in (c) stem from the reduced efficiency due to the RF foil in the VELO.

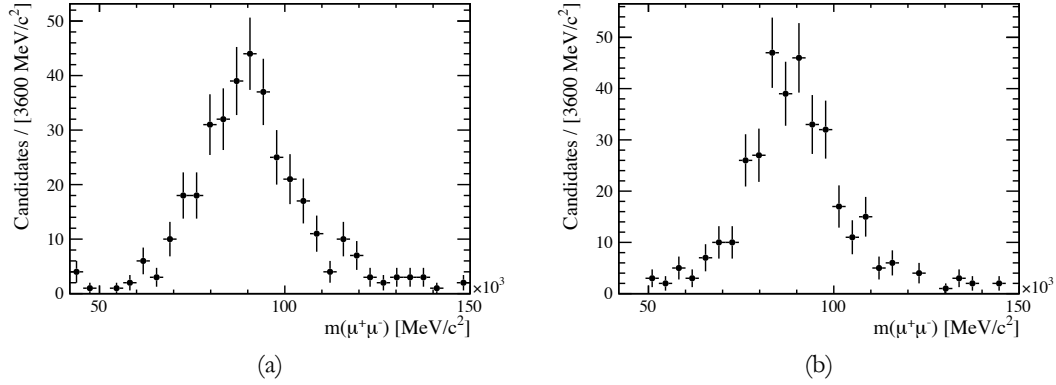


Figure 4.10.: Dimuon invariant mass distribution for muons made of a long track and a MuonTT track in (a) with  $\mu^-$  as probe-track and in (b) with  $\mu^+$  as probe-track. The selection cuts as written in Table 4.5 are applied. The distribution peaks close to the nominal position of the  $Z^0$  mass.

Reconstructible in this context means that the simulated particle left enough information<sup>17</sup> in the subdetectors to make a reconstruction possible. This is called «hit-based method» in the following. The results of this comparison can be seen in Fig. 4.9. Both efficiencies agree within  $\approx 1\%$  in each bin, which is well below the precision which can be reached with the tag-and-probe method on 2010 collision data.

The uncertainty is calculated using the Bayesian method described in Ref. [111]. As the sample with the associated tracks is a strict subsample of the sample with all tracks, a binomial error would normally be chosen. However, for efficiencies close to 100% the uncertainty would become very small and would completely vanish for efficiencies of 100%. This was not deemed appropriate as sometimes very small statistics are present in the individual bins. Therefore the alternative Bayesian approach is used.

### 4.5.3. Checks on collision data

Two major crosschecks are performed on collision data: The comparison between the tracking efficiency taking the negative and the positive muon as probe, respectively, and checking for differences between data taking with upward or downward magnetic field.

**Consistency between both charges of the muon** The result of the first check can be seen in Fig. 4.13. No significant discrepancy between the two charges can be observed. The overall numbers are given in Table 4.6.

<sup>17</sup>Three hits in the VELO  $r$ -sensors, three hits in the VELO  $\phi$ -sensors, one hit in an  $x$ -plane in each of the T-stations and 1 hit in a stereo layer in each of the T-stations.

**Consistency between both magnet polarities** The result of the check for differences in efficiency for different magnetic field directions is shown in Fig. 4.14 and the overall numbers are given in Table 4.6. All these numbers are compatible within the statistical uncertainty.

Table 4.6.: Consistency checks for the tracking efficiency measurement for high transverse momentum tracks for different charges of the probe-muon and different directions of the magnetic field.

Magnetic field	Charge of probe	Efficiency [%]
up & down	positive	$96.6^{+0.9}_{-1.2}$
	negative	$96.4^{+1.0}_{-1.2}$
up	positive & negative	$96.8^{+1.0}_{-1.2}$
down		$96.4^{+0.9}_{-1.1}$

#### 4.5.4. Results

The tracking efficiency measured using the  $37.5 \text{ pb}^{-1}$  of collision data recorded in 2010 are shown in Fig. 4.11. The distribution of  $Z^0 \rightarrow \mu^+ \mu^-$  events where the negative muon is reconstructed as a MuonTT track is shown in Fig. 4.12. Overall the tracking efficiency for these high transverse momentum tracks is between 95% and 100% with an uncertainty of  $\pm 2\text{-}3\%$ . It is mostly constant as a function of pseudorapidity, transverse momentum and polar angle. It decreases as a function of the number of primary vertices, although the experimental precision is not good enough to make a clear statement. The average tracking efficiency, integrated over all bins, is  $(96.5 \pm 0.8)\%$ . No systematic effects are considered for this determination given the large statistical uncertainties.

The tracking efficiency method presented in this section was used for the measurements of the  $W$  and  $Z^0$  cross-sections [110]. The tracking efficiency measurement for high transverse momentum tracks was not repeated for 2011 collision data. As the statistical uncertainty would decrease by a large amount for these data due to the much larger data set, a study of the background contribution and the assignment of a systematical uncertainty would certainly be necessary.

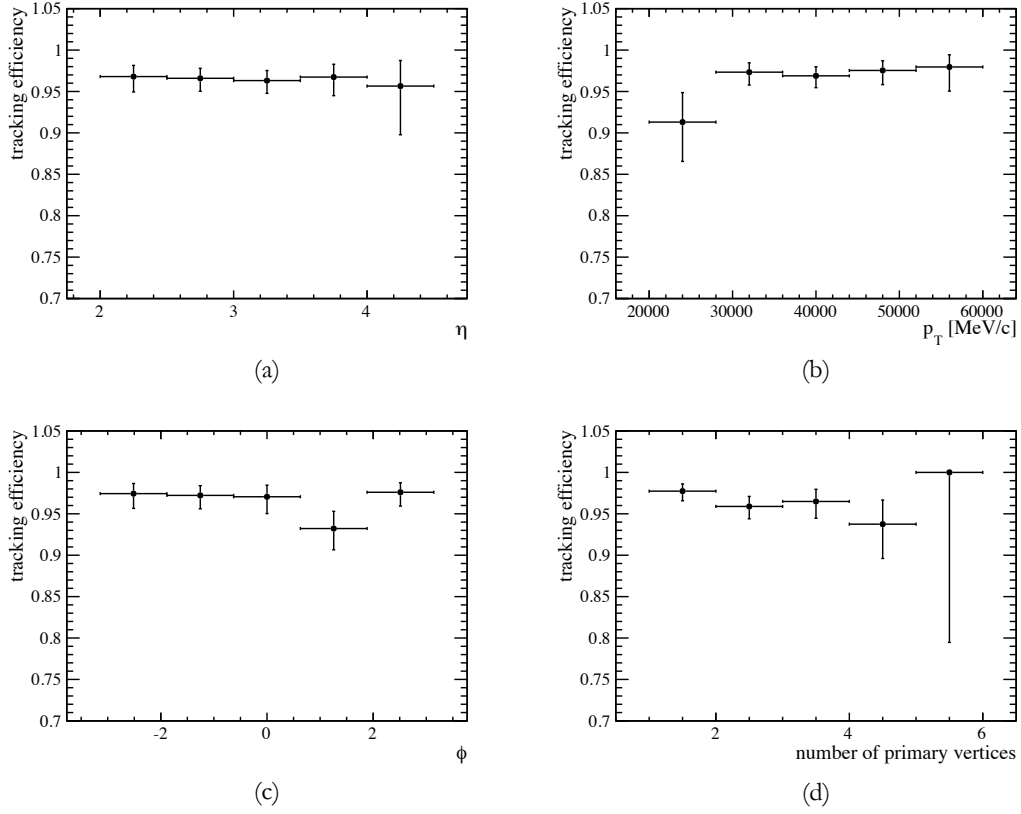


Figure 4.11.: Tracking efficiency for high transverse momentum tracks, determined in 2010 collision data with tag-and-probe on  $Z^0 \rightarrow \mu^+ \mu^-$ . The efficiency is binned in the pseudorapidity (a), the transverse momentum (b), the polar angle (c) and the number of reconstructed primary vertices (d). The uncertainty is calculated using the Bayesian technique described in Ref. [111].

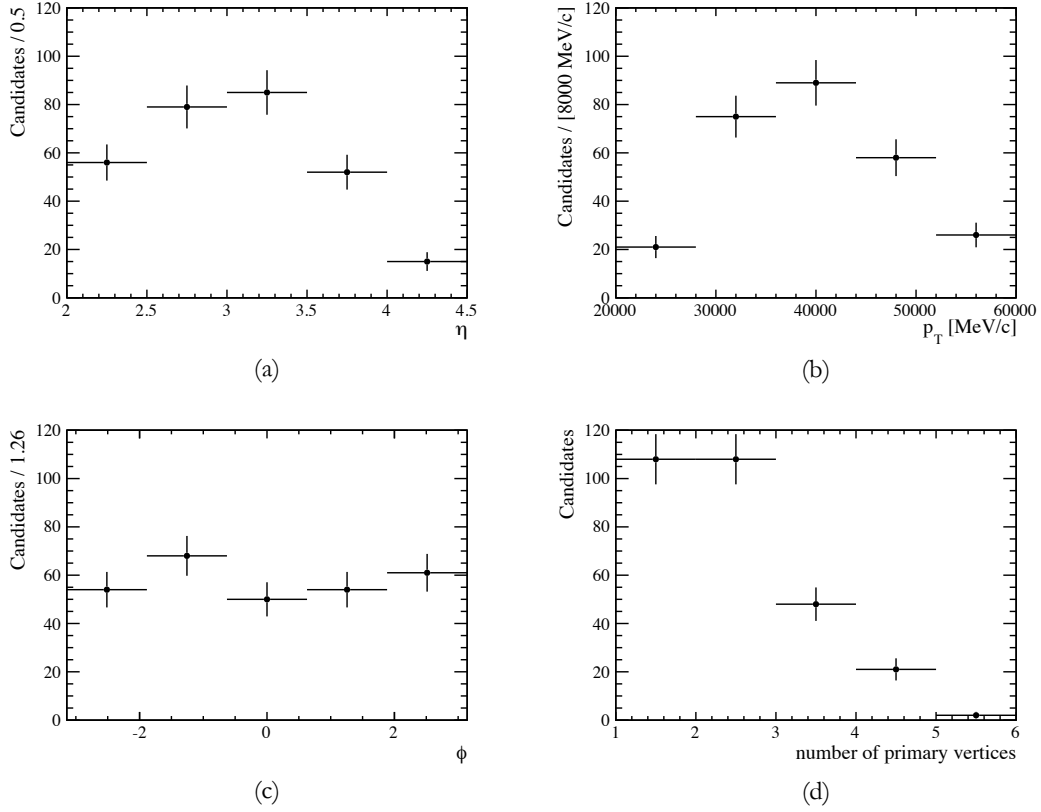


Figure 4.12.: Distribution of  $Z^0 \rightarrow \mu^+ \mu^-$  events, where the negative muon was reconstructed as a MuonTT track, as a function of the pseudorapidity of the MuonTT track (a), transverse momentum of the MuonTT track (b),  $\phi$ -angle of the MuonTT track (c) and number of primary vertices (d) in the event.

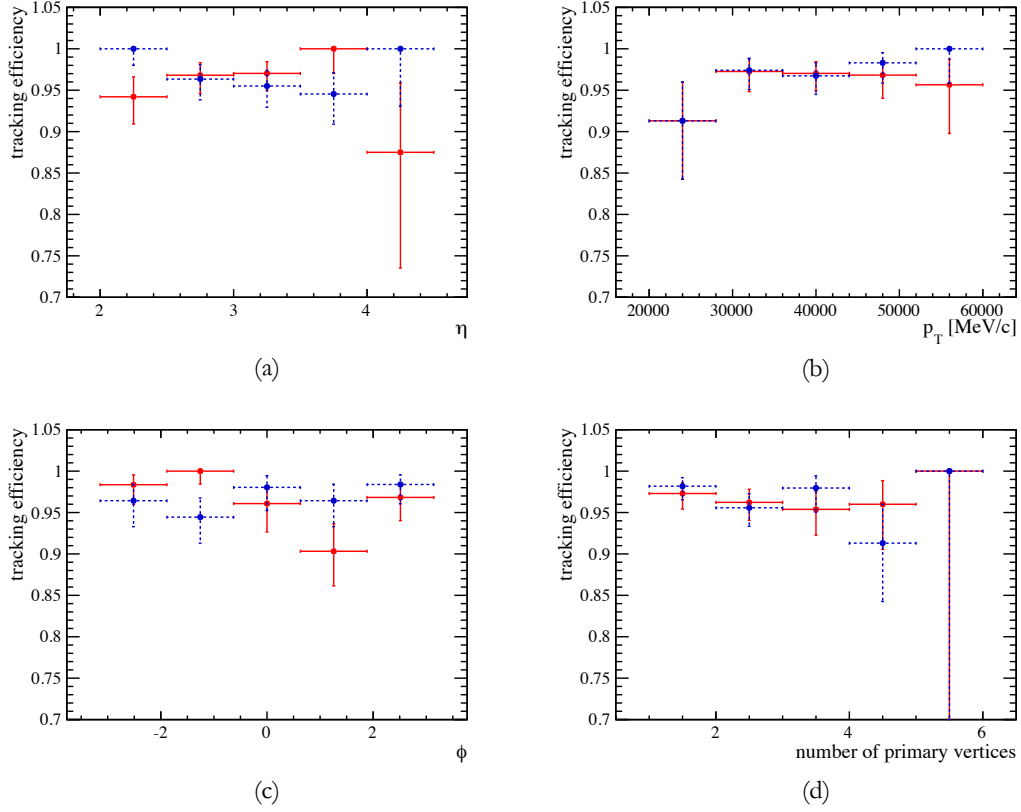


Figure 4.13.: Tracking efficiency for high transverse momentum tracks, determined in 2010 collision data with tag-and-probe on  $Z^0 \rightarrow \mu^+ \mu^-$ . The efficiency is binned in the pseudorapidity (a), the transverse momentum (b), the polar angle (c) and the number of reconstructed primary vertices (d). The red lines describe the efficiency for positive muons, the blue dashed line the efficiency for negative muons.

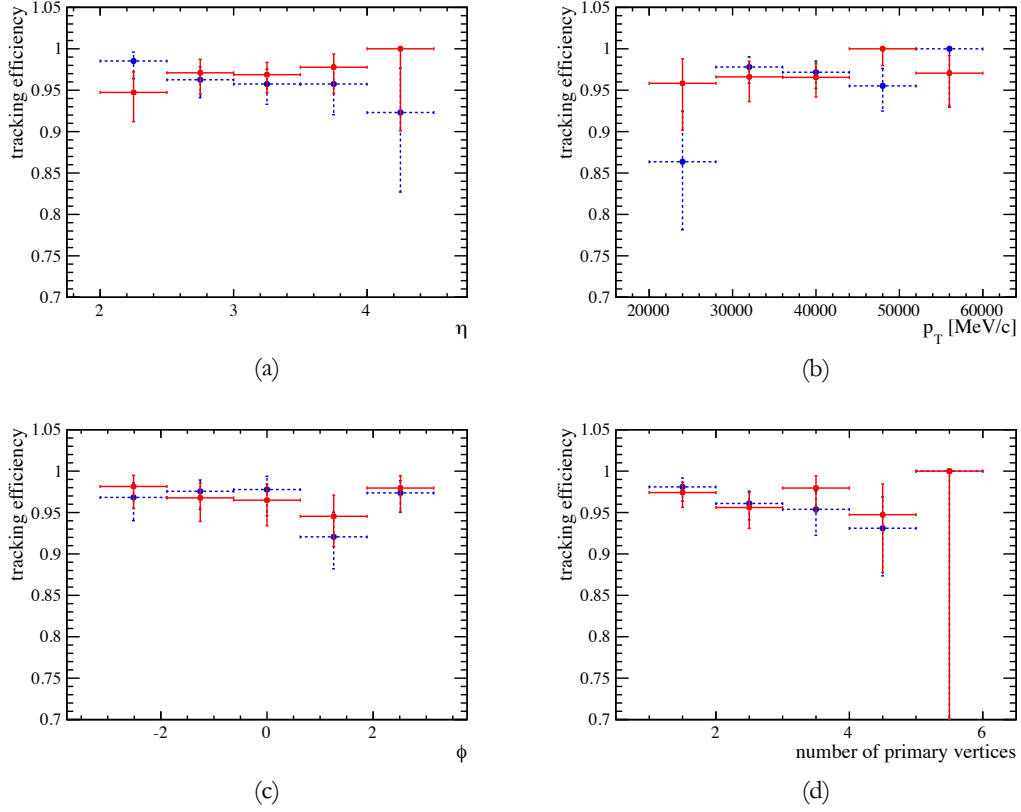


Figure 4.14.: Tracking efficiency for high transverse momentum tracks, determined in 2010 collision data with tag-and-probe on  $Z^0 \rightarrow \mu^+ \mu^-$ . The efficiency is binned in the pseudorapidity (a), the transverse momentum (b), the polar angle (c) and the number of reconstructed primary vertices (d). The red lines describe the efficiency for upward magnetic field, the blue dashed line the efficiency for downward magnetic field.

## 4.6. Tracking efficiency for low transverse momentum tracks

Most of the decay channels studied in LHCb deal with tracks of low or moderate (transverse) momentum, *i.e.* decay products of  $B$  or  $D$  meson decays. Therefore it is essential to use a dimuon resonance to determine the tracking efficiency which provides approximately the same momentum spectrum. The decay  $J/\psi \rightarrow \mu^+\mu^-$  is an ideal candidate as it is abundant and its peak very narrow. However, neither in 2010 nor in 2011 there was no single muon trigger line in HLT2 during 2010 and 2011 which would have allowed the determination of the tracking efficiency with prompt  $J/\psi$ <sup>18</sup>. Therefore, only non-prompt  $J/\psi$ , mostly from  $B$  decays, could be used.

As the MuonTT tracking algorithm expects the muon to come from the primary vertex, the restriction to non-prompt  $J/\psi$  was not ideal. However, it turned out that a large amount of  $J/\psi$  can still be reconstructed due to the large search windows in the pattern recognition in TT, and the efficient removal of hits which do not correspond to the MuonTT track.

The phase-space distributions in pseudorapidity and momentum for MuonTT tracks from non-prompt  $J/\psi$  and for long tracks from  $B^0 \rightarrow J/\psi K^{*0}$  is shown in Fig. 4.15. The  $B^0 \rightarrow J/\psi K^{*0}$  decay serves as a comparison channel as the momentum and pseudorapidity of its decay products are typical for many decay channels LHCb is investigating. It can be seen that the phase-space covered by the MuonTT tracks from  $J/\psi$  largely covers the one for the  $B^0 \rightarrow J/\psi K^{*0}$  decay products.

For the tracking efficiency measurement for low transverse momentum tracks, a different strategy than for  $Z^0 \rightarrow \mu^+\mu^-$  was used. The efficiency was binned in two bins in pseudorapidity  $\eta$  and five bins in momentum  $p$  to make the results more easily usable for a wider range of decays. Furthermore, the final results were given as ratios between the tracking efficiency determined on collision data and on simulated data. This approach was chosen to reduce the amount of systematic uncertainties. Finally, the results from the Combined Method and the Long Method were joined in one ratio-table with five bins in momentum and two in pseudorapidity.

### 4.6.1. Event selection

In the stripping some cuts are applied to the  $J/\psi$  candidate in order to reduce the rate and the combinatorial background. These selection cuts are summarised in Table 4.9. Note that certain new cuts were introduced from 2010 to 2011 while others were tightened. Moreover, for the 2011 stripping selection, the cuts on the tag-muon were aligned with the cuts for the Hlt2SingleMuon trigger line, which was not the case in 2010. The offline selection is summarised in Table 4.10. To reject spurious tracks with a high chance of being

---

<sup>18</sup>In the early runs of 2010 there was a trigger line which could have been used. However, these data were never stripped for tracking efficiency studies.



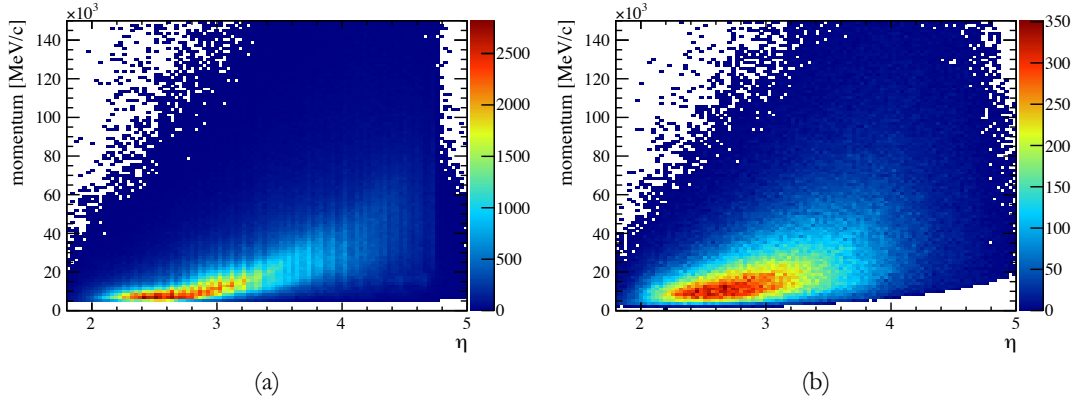


Figure 4.15.: Distribution of tracks in pseudorapidity and momentum for MuonTT tracks from  $J/\psi$  (a) and tracks coming from  $B^0 \rightarrow J/\psi K^{*0}$  (b).

background, cuts were applied on the probe-muon. The effect of the offline selection alone is shown in Fig. 4.17.

The selection criteria used was very similar for 2010 and 2011 collision data. Differences are only present in variables which do not bias the result.

#### 4.6.2. Data samples

The tracking efficiency with the Long Method was determined on two different samples: The full 2010 data sample and the full 2011 data sample. The integrated luminosities of these samples are summarised in Table 4.7.

Table 4.7.: Integrated luminosity used for determinations of tracking efficiency on 2010 and 2011 collision data.

Data type	Integrated luminosity $\text{pb}^{-1}$
2010	29 <sup>19</sup>
2011	1027

<sup>19</sup>The number for 2010 data is smaller than the corresponding one for  $Z^0 \rightarrow \mu^+ \mu^-$ . This has purely technical reasons in the stripping as not all data could be stripped for the tracking efficiency measurement in 2010.

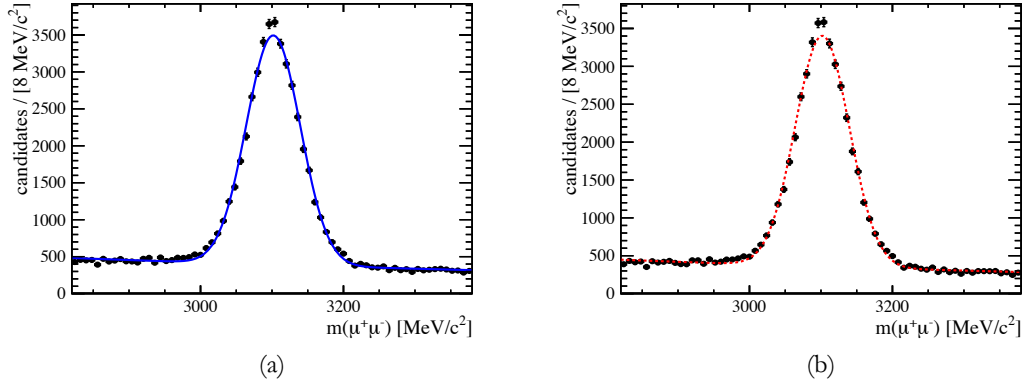


Figure 4.16.: Dimuon invariant mass distribution for all long track – MuonTT track combinations (a), and long track – MuonTT track combinations where the MuonTT track could be matched to a long track (b). Shown are events with the probe-track in the bin corresponding to  $20 \text{ GeV}/c < p < 40 \text{ GeV}/c$  and  $1.9 < \eta < 3.2$ .

#### 4.6.3. Fit model

The tracking efficiency for  $J/\psi \rightarrow \mu^+\mu^-$  was extracted by fitting the dimuon mass spectrum for all long track – MuonTT track combinations and for only those combinations where the MuonTT track could be associated to a long track with a  $\chi^2/\text{ndof} < 5$ . This criterion ensures that no tracks are matched that are ghosts or otherwise very poorly reconstructed (see Sect. 4.6.7 for a further discussion of this cut). The model for the mass distribution is a single Gaussian and an exponential background. An example is shown in Fig. 4.16 for the range  $20 \text{ GeV}/c < p < 40 \text{ GeV}/c$  and  $1.9 < \eta < 3.2$ . The width of the Gaussian was fixed when fitting the associated events to the value obtained when fitting all events. Although a more complicated fit model would in principle be appropriate, *e.g.* a double Crystal Ball function, these fits were very unstable and showed problems converging or a tendency for certain parameters to have unphysical values for  $p$ - $\eta$  bins with low statistics. In these cases, the tail of the double Crystal Ball incorrectly incorporates parts of the background, which will yield a wrong efficiency or will at least make the efficiency difficult to interpret. To avoid these complications and have a consistent mass model for all bins, a single Gaussian solution was chosen. As only ratios of event yields are considered, the error from using a single Gaussian only largely cancels.

#### 4.6.4. Checks on simulated data

**Efficiency and purity for association** The association of the MuonTT track to the long track is based on a comparison between hit distributions for long tracks that do and do not

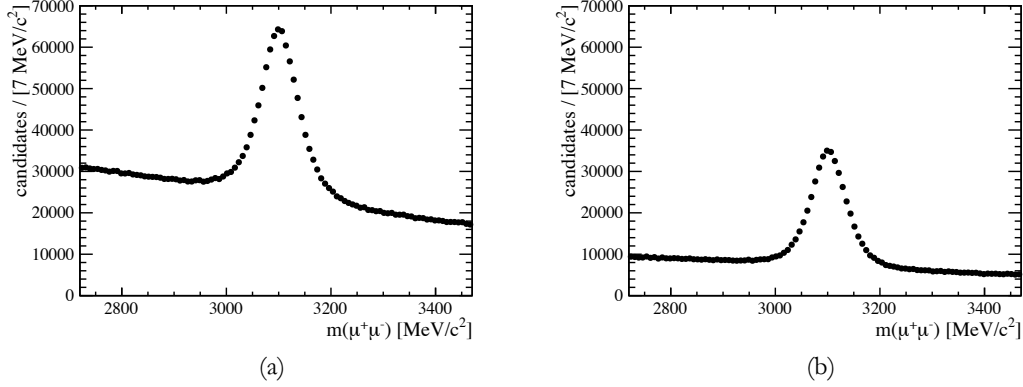


Figure 4.17.: Invariant mass distribution of dimuon pairs for the tracking efficiency determination using  $J/\psi$  after the stripping selection (a); and after the stripping and offline selection (b).

correspond to a MuonTT track. The efficiency and purity of this association is determined on simulated  $J/\psi \rightarrow \mu^+ \mu^-$  events, where it is required that both muons are reconstructed as long tracks and can be matched to the corresponding simulated particles. It is also required that the muons are stemming from a  $J/\psi$ . Furthermore a MuonTT track needs to be reconstructed. In this sample, the efficiency is therefore 100% by definition. The matching efficiency is determined by requiring a track to pass the association criterion; it is 99.5%, with a negligible uncertainty. The purity is 99.9%, and also has a negligible uncertainty. The purity is determined in a similar way, by requiring that the associated track is identical to the track coming from the  $J/\psi$ . As both values are very close to one, and the effect cancels when building the tracking efficiency ratio in collision data and simulation, this is not considered any further.

**Hit-based tracking efficiency in simulation** As for  $Z^0 \rightarrow \mu^+ \mu^-$  decays, the presence of systematic errors in the tracking efficiency determination using  $J/\psi \rightarrow \mu^+ \mu^-$  decays was checked. This is again done by comparing the numbers obtained in simulation with the tag-and-probe method, and directly from simulation with the hit-based method. The efficiency is determined with the same binning scheme as used in collision data. The projections are shown in Fig. 4.18 and the agreement is of the order of 1% for each bin.

As for the efficiency and purity check, this effect cancels when building the tracking efficiency ratio between collision data and simulation.

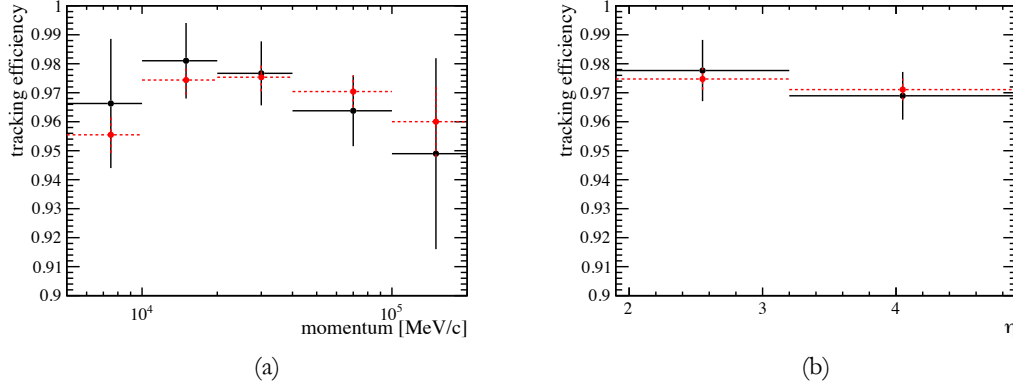


Figure 4.18.: Comparison between the tracking efficiency determined on simulation level with the hit-based method (red dashed data points) and with the tag-and-probe approach (black solid data points) on a sample of simulated  $J/\psi \rightarrow \mu^+\mu^-$ . (a) shows the projection in momentum and (b) the projection in pseudorapidity. Both methods agree within  $\approx 1\%$ .

#### 4.6.5. Checks on collision data

To check possible systematic effects stemming from the different magnet polarities and the different charge of the probe-muon, the 2010 and the 2011 collision data samples were split in four subsamples, each corresponding to a specific magnet polarity and probe-muon charge. The tracking efficiency was determined for the full momentum and pseudorapidity range with a fit to all long track and MuonTT track combinations and only those where the MuonTT track can be matched to a long track. The results of these comparisons are shown in Table 4.8. The efficiencies for the 2011 sample agree very well with each other. However the subsamples with upwards magnetic field in the 2010 data tend to have a lower efficiency than the ones with a downward magnetic field.

#### 4.6.6. Reweighting

The tracking efficiency mainly depends on three variables: the momentum, the pseudorapidity and the number of tracks per event. In principle the efficiency should then be determined in bins of all three variables. However, given the low statistics, this is not feasible. A simplified approach was therefore chosen: In order to have the same number of tracks per event in simulation and collision data, the simulated events are reweighted with respect to number of tracks per event to the corresponding value from collision data. This approach was chosen as this number is constant for almost all of the physics analyses. To avoid large uncertainties on the reweighting factor, only weights with a relative uncertainty of less than 30% are applied, otherwise the weight is set to one. This introduces a negli-

gible bias. The distribution of the number of tracks per event for simulation, simulation after reweighting, and data is shown in Fig. 4.20. The tracking efficiency is given as the ratio between collision data and reweighted simulation data in the remaining two variables, momentum and pseudorapidity. The comparison between simulation and collision data for all three variables for the full 2010 and 2011 data sample is shown in Fig. 4.19.

#### 4.6.7. Impact of track quality cut

As mentioned in Sect. 4.6.3, only MuonTT tracks that can be matched to long tracks with a  $\chi^2_{\text{track}}/\text{ndof} < 5$  are considered associated. However, the  $\chi^2$  distribution of tracks in collision data and in simulated data is not identical, and therefore, the ratio of tracking efficiencies in collision data and in simulation could depend on the value of this cut. In Fig. 4.21, it is shown that this is not the case as the ratio stays constant above a cut of  $\chi^2_{\text{track}}/\text{ndof} < 4$ .

#### 4.6.8. Results for the Long Method

The results for the Long Method are presented as two-dimensional maps for the efficiency in collision data and the ratio between the efficiency in collision data and in simulation, see Fig. 4.22. For almost all bins the ratio is close to one, meaning that the tracking efficiency in simulation is a good approximation for the tracking efficiency in collision data. Note that the two bins for low momentum / high pseudorapidity and high momentum / low pseudorapidity show large deviations from one. The amount of statistics used to determine the efficiency in these bins is very low (also compare this to the distribution in Fig. 4.15) and therefore no precise statement about the tracking efficiency can be made. Most of the analyses in LHCb do not have a large fraction of their signal events in this region so the impact on the final precision is not substantial.

#### 4.6.9. Calculation of uncertainties

Two types of uncertainties need to be calculated: one on the efficiency and one on the efficiency ratio between simulation and collision data.

**Uncertainty on the efficiency** The uncertainty on the efficiency consists of three parts: an uncertainty on the number of associated tracks (corresponding to the signal events); an uncertainty on the number of background events; and an uncertainty on the number of associated background events. The total number of events does not have an uncertainty by construction.

The uncertainty on the efficiency can be calculated using only the total and associated number of signal events and their respective uncertainties. However a different approach was

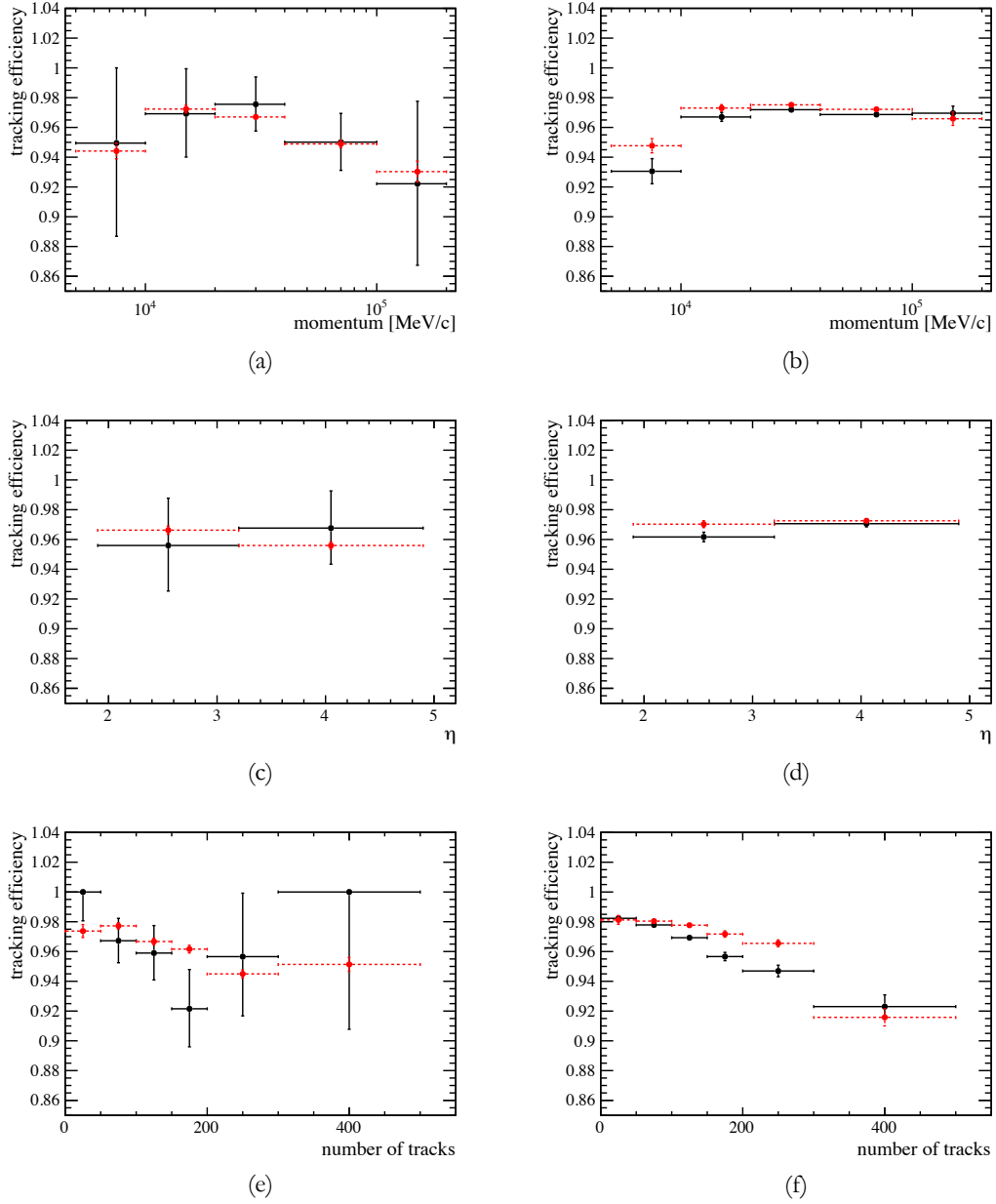


Figure 4.19.: Comparison between the tracking efficiency in the simulation (red dashed data points) and in collision data (black solid data points) with the Long Method in the momentum ((a), (b)), the pseudorapidity ((c), (d)) and the number of tracks ((e), (f)). The left column corresponds to 2010 collision data, the right to 2011 collision data.

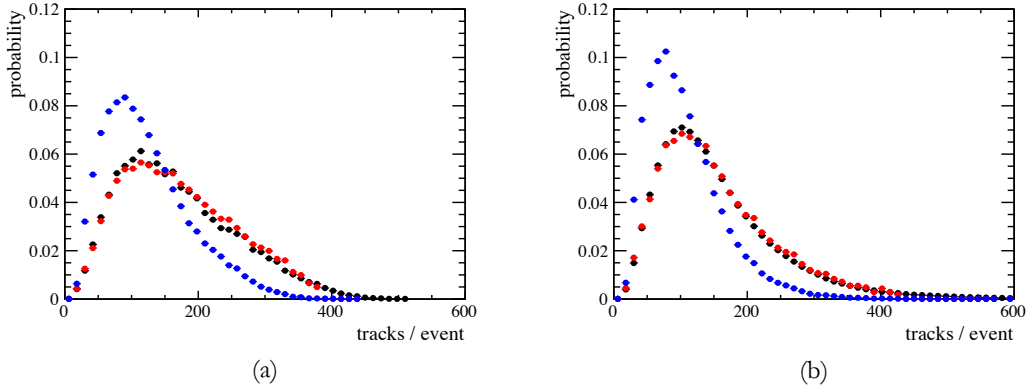


Figure 4.20.: Comparison between (normalised) distribution of tracks per event for the simulated sample (blue), the reweighted simulated sample (red) and the data sample (black) for 2010 collision data (a) and 2011 collision data (b). The drop-off at about 400 tracks per event in the reweighted simulated samples is due to the cut on the relative uncertainty of the weight.

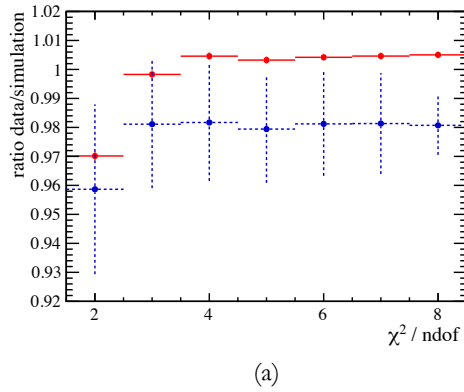


Figure 4.21.: Ratio between tracking efficiency on collision data and simulation as a function of the applied  $\chi^2/\text{ndof}$  cut on the track associated to the MuonTT track. The red solid line corresponds to the 2011 data while the dashed blue line is the 2010 data.

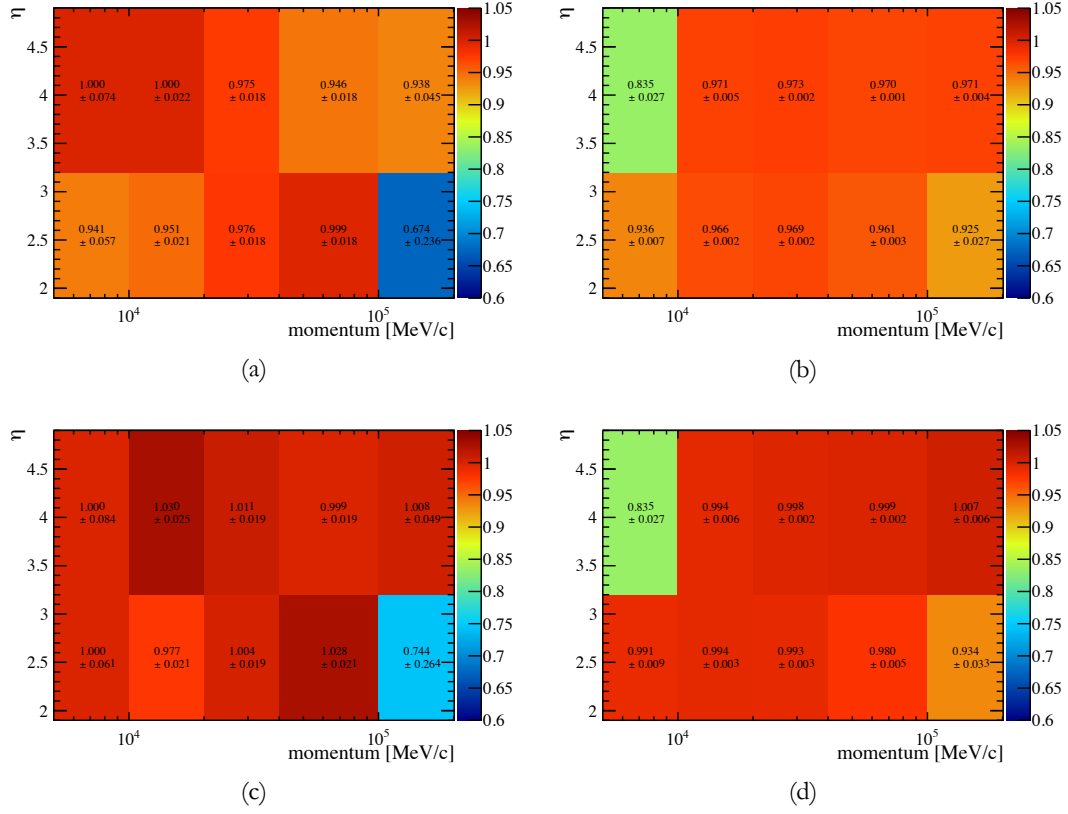


Figure 4.22.: Tracking efficiency measured using the Long Method as a function of momentum and pseudorapidity for 2010 collision data (a) and 2011 collision data (b). Figures (c) and (d) show the ratio between the tracking efficiency in collision data and in (reweighted) simulation for 2010 (c) and 2011 (d) with the Long Method. The two bins for high  $\eta$  and low momentum, low  $\eta$  and high momentum, respectively, suffer from small statistics and therefore will not correctly reproduce the tracking efficiency ratio.



chosen. As the Combined Method (see Sect. 4.7) uses a «simple sideband subtraction»<sup>20</sup> to determine the number of signal events, a method was developed for the Long Method to adapt the same procedure, starting from results of the fit to the invariant mass distribution<sup>21</sup>. In the sideband subtraction, knowing the number of events in the signal window  $n_{\text{tot}}$  (containing background and signal) and in the sidebands  $n_{\text{bkg}}$  (containing background only) allows the calculation of all three uncertainties needed for the tracking efficiency. The uncertainty on the signal can be used to derive  $n_{\text{tot}}$ . The invariant mass range is by definition twice the size of the signal window in a sideband subtraction and can be written as:

$$\sigma_{\text{sig}} = \sqrt{n_{\text{sig}} + 2 \cdot n_{\text{bkg}}}. \quad (4.6)$$

The total number of events in the signal window is then  $n_{\text{tot}} = \frac{1}{2}(\sigma_{\text{sig}}^2 + n_{\text{sig}})$ . Furthermore the number of background events is  $n_{\text{bkg}} = n_{\text{tot}} - n_{\text{sig}}$ .

Starting from the number of (associated) signal and background events, the three uncertainties were then calculated. The Feldmann-Cousins method [112] with a 68% confidence level was used if the corresponding numbers were small and the values were close to the boundaries. The binomial error (with a 68% confidence level based on a Gaussian distribution) was used otherwise:

$$\sigma_{\text{eff, binom.}} = \frac{\sqrt{n_{\text{sig, assoc}} \left(1 - \frac{n_{\text{sig, assoc}}}{n_{\text{sig, tot}}}\right)}}{n_{\text{sig, tot}}}, \quad (4.7)$$

where  $n_{\text{tot}}$  is the total number of events and  $n_{\text{assoc}}$  is the number of events where a MuonTT track can be matched to a long track. The binomial uncertainty for the background is defined likewise. Note that the Feldmann-Cousins method returns asymmetric uncertainties, while the binomial uncertainty is symmetric by construction. The total upper and lower uncertainty can then be calculated as:

$$\sigma_{\text{eff, up}} = \sqrt{\sigma_{\text{sig, assoc, up}}^2 + \sigma_{\text{bkg, tot, up}}^2 + \sigma_{\text{bkg, assoc, low}}^2} \quad (4.8)$$

$$\sigma_{\text{eff, low}} = \sqrt{\sigma_{\text{sig, assoc, low}}^2 + \sigma_{\text{bkg, tot, low}}^2 + \sigma_{\text{bkg, assoc, up}}^2}. \quad (4.9)$$

where  $\sigma_{\text{sig, assoc / tot, up / down}}$  denote the uncertainties stemming from the upper / lower uncertainty on the total / associated number of signal events and  $\sigma_{\text{bkg, assoc / tot, up / down}}$  the corresponding uncertainties stemming from the uncertainties for the background events.

<sup>20</sup>In an invariant mass distribution with a signal peak and a linear background, the background can be statistically subtracted by determining the amount of background in a sideband left and right of the peak, where each sideband has half the size of the signal window. The number of signal events is then the total number of events in the signal window minus the background events in both sidebands.

<sup>21</sup>The methods are mathematically equivalent.

The uncertainty is constructed such that it takes into account the total possible fluctuation for the number of associated and total background events.

**Uncertainty on the efficiency ratio** The uncertainty on the ratio  $r$  between the efficiency in collision data and in simulation is calculated using the formula:

$$\sigma_{r, \text{up}} = \sqrt{\left(\frac{\sigma_{\text{data, eff, up}}}{\varepsilon_{\text{sim}}}\right)^2 + \left(\frac{r \cdot \sigma_{\text{sim, eff, low}}}{\varepsilon_{\text{sim}}}\right)^2} \quad (4.10)$$

$$\sigma_{r, \text{low}} = \sqrt{\left(\frac{\sigma_{\text{data, eff, low}}}{\varepsilon_{\text{sim}}}\right)^2 + \left(\frac{r \cdot \sigma_{\text{sim, eff, up}}}{\varepsilon_{\text{sim}}}\right)^2} \quad (4.11)$$

An average uncertainty is then calculated using these results and assigned as a statistical uncertainty to the tracking efficiency measurement made with the Long Method.

Table 4.8.: Consistency checks for the tracking efficiency measurement for low transverse momentum tracks for different charges of the probe-muon and different directions of the magnetic field.

Data Type	Magnetic field	Charge of probe	Efficiency [%]
2010 data	up	positive	$93.2^{+1.5}_{-1.7}$
		negative	$94.4 \pm 1.5$
	down	positive	$97.8 \pm 2.0$
		negative	$98.7^{+1.3}_{-1.7}$
2011 data	up	positive	$97.0 \pm 0.1$
		negative	$97.0 \pm 0.2$
	down	positive	$96.8 \pm 0.1$
		negative	$96.8 \pm 0.1$

Table 4.9.: Selection cuts applied on the  $J/\psi$  resonance in the stripping for the tracking efficiency measurement on 2010 and 2011 data. The tighter cuts in 2011 data are due to tighter requirements in the stripping, as stricter timing and rate conditions were imposed.

Variable	Value (2010)	Value (2011)
$p$ tag-muon	–	$> 10'000 \text{ MeV}/c$
$p_T$ tag-muon	$> 1000 \text{ MeV}/c$	$> 1300 \text{ MeV}/c$
$\chi^2_{\text{track}}/\text{ndof}$ tag-muon	$< 5$	$< 5$
isMuon tag-muon	true	true
$\text{DLL}_{\mu\pi}$ tag-muon	$> 0$	$> 2$
$\text{IP}_{\min}$ tag-muon <sup>22</sup>	–	$> 0.5 \text{ mm}$
$\chi^2_{\text{IP}}$	–	$> 200$
$J/\psi \chi^2_{\text{vtx}}/\text{ndof}$	$< 10$	$< 5$
$J/\psi$ mass window	$500 \text{ MeV}/c^2$	$500 \text{ MeV}/c^2$
$J/\psi p_T$	$> 500 \text{ MeV}/c$	$> 1000 \text{ MeV}/c$
$J/\psi \text{IP}_{\min}$	–	$< 3 \text{ mm}$
Prescale	1	0.5

## 4.7. Alternative method to measure the tracking efficiency for long tracks

Instead of constructing a MuonTT track to measure the tracking efficiency with the tag-and-probe approach in one go, the measurement can also be split up in two separate parts: a measurement of the VELO tracking efficiency (VELO Method) and a measurement of the T station tracking efficiency (T Station Method). Both methods use tag-and-probe separately. By combining these two results, the efficiency to reconstruct a long track can be calculated. This is known as the Combined Method.

### 4.7.1. VELO tracking efficiency

To determine the tracking efficiency of the VELO, one can use probe-tracks with segments in the TT and the T stations, called downstream tracks (see Sect. 3.4). However, in the standard track reconstruction, the downstream track is not stored on disk if a long track is present that shares more than 70% of the hits in TT and the T stations with the downstream

<sup>22</sup> $\text{IP}_{\min}$  is the smallest impact parameter with respect to all reconstructed primary vertices in the event.

Table 4.10.: Offline selection cuts applied for the  $J/\psi$  resonance for the tracking efficiency measurement on 2010 and 2011 data. The cuts on the probe-muon are for the rejection of spurious tracks which would have a high probability of being background. Only additional or tighter cuts with respect to Table 4.9 are mentioned.

Variable	Value (2010)	Value (2011)
$p$ tag-muon	$> 10'000 \text{ MeV}/c$	$> 10'000 \text{ MeV}/c$
$p_T$ tag-muon	$> 1300 \text{ MeV}/c$	$> 1300 \text{ MeV}/c$
$\chi_{\text{track}}^2/\text{ndof}$ tag-muon	$< 4$	$< 4$
$p$ probe-muon	$> 5000 \text{ MeV}/c$	$> 5000 \text{ MeV}/c$
$p_T$ probe-muon	$> 100 \text{ MeV}/c$	$> 100 \text{ MeV}/c$
$J/\psi \chi_{\text{vtx}}^2/\text{ndof}$	$< 5$	$< 5$
$J/\psi \text{ IP}_{\text{min}}$	$< 0.8 \text{ mm}$	$< 0.8 \text{ mm}$
# $J/\psi$ candidates/event	1	1

track. Therefore the downstream track finding needs to be redone to obtain the full sample of probe-tracks. As for the Long Method the tag-track needs to have caused the trigger decision in HLT1 and HLT2 in order not to be biased by the trigger. Furthermore a selection in the stripping and offline is applied to select  $J/\psi \rightarrow \mu^+\mu^-$  decays (more details can be found in Ref. [100]), where the number of reconstructed  $J/\psi$  is determined using a simple sideband subtraction. The downstream track is considered associated to a long track if both tracks share more than 50% of their hits in TT and the T stations.

#### 4.7.2. T station tracking efficiency

Instead of excluding the VELO in the long track finding, one can also exclude the T stations. The calorimeters or the muon system then serve as a tracking station downstream of the magnet. While the first one is used in efficiency determinations with  $K_S^0 \rightarrow \pi^+\pi^-$ , the latter is used with  $J/\psi \rightarrow \mu^+\mu^-$  decays. Both algorithms are explained in Ref. [104], but only the one for  $J/\psi \rightarrow \mu^+\mu^-$  will be described here (as this one allows the combination with the VELO Method and the comparison with the Long Method). The muon tracks are found by an algorithm similar to the one used to construct MuonTT tracks (see Ref. [113] for the original algorithm), while the VELO tracks are found by the standard VELO pattern recognition. The VELO and the muon segment are then extrapolated to the central plane of the magnetic field to measure their distance in  $x$  and to M2 to measure their distance in  $y$ . The latter position was chosen as the extrapolation error for VELO tracks due to multiple scattering is smaller than the one for muon tracks. The compatibility

between VELO and muon segment is then checked using a weighted quadratic sum for the individual distances in the  $x$ - and  $y$  directions. As for downstream tracks in the VELO Method, the VELO pattern recognition has to be redone, as these tracks are not kept when they serve as a seed for a long track.

The long track and the VELO-muon track are then combined to perform the tag-and-probe approach in  $J/\psi \rightarrow \mu^+ \mu^-$  decays with similar selections and trigger requirements as for the VELO Method, where a simple sideband subtraction is again used to determine the number of  $J/\psi$ . The VELO-muon track is considered associated to a long track if both tracks share the same hits in the VELO and at least two hits in the muon stations. The VELO pattern recognition prevents the construction of two VELO tracks which share more than 70% of the hits. Furthermore, a long track has the same VELO hits as its VELO seed: both properties ensure that via the matching condition in the VELO, the same VELO track segment is used for the long track and the VELO-muon track.

## 4.8. Combined results for low transverse momentum tracks

The efficiencies obtained with the VELO Method and the T Station Method are multiplied together to obtain the efficiency to reconstruct long tracks (Combined Method). This efficiency can then be compared with the result from the Long Method. Although both efficiencies are constructed in a similar way, a perfect agreement is not expected as the Long Method and the VELO Method for example measure by construction the efficiency for tracks in the acceptance of the TT, while the T Station Method does not. Using the ratio of efficiencies in collision data and simulation, however, largely cancels these effects. The efficiencies from the Long and the Combined Method are combined using the weighted average in each pseudorapidity and momentum bin,  $r_{\text{long track}}$ :

$$r_{\text{long track}} = \frac{r_{\text{Long Method}} \cdot \sigma_{\text{Combined Method}}^2 + r_{\text{Combined Method}} \cdot \sigma_{\text{Long Method}}^2}{\sigma_{\text{Long Method}}^2 + \sigma_{\text{Long Method}}^2}, \quad (4.12)$$

with  $\sigma_{\text{Long Method}}$  and  $\sigma_{\text{Combined Method}}$  the uncertainty on the ratio between collision data and simulation for the Long Method and the Combined Method respectively.

The statistical uncertainty on the tracking efficiency ratio can then be calculated using standard error propagation:

$$\sigma_{r_{\text{long track, stat}}} = \frac{\sigma_{\text{Long Method}} \cdot \sigma_{\text{Combined Method}}}{\sqrt{\sigma_{\text{Long Method}}^2 + \sigma_{\text{Combined Method}}^2}}. \quad (4.13)$$

Furthermore, the difference of the individual methods with respect to their weighted average, not already being taken into account in their respective statistical uncertainties, is added as a systematic uncertainty<sup>23</sup>:

$$\sigma_{r_{\text{long track, sys}}}^2 = \left(r_{\text{Long Method}} - r_{\text{long track}}\right)^2 + \left(r_{\text{Combined Method}} - r_{\text{long track}}\right)^2 - \sigma_{\text{Long Method}}^2 - \sigma_{\text{Combined Method}}^2 \quad (4.14)$$

The total uncertainty is given as the quadratic sum of both uncertainties:

$$\sigma_{r_{\text{long track}}} = \sqrt{\sigma_{r_{\text{long track, stat}}}^2 + \sigma_{r_{\text{long track, sys}}}^2} \quad (4.15)$$

#### 4.8.1. Additional systematic effects

The reweighting of the simulation sample is done in number of tracks, number of primary vertices or number of hits in different subdetectors to evaluate systematic effects. The largest difference arising in this procedure is 0.4%, which is taken as the systematic uncertainty of the final ratio. The total uncertainty is that given by Eq. 4.15 plus 0.4%, added in quadrature.

#### 4.8.2. Results

The results from the combination of the Long Method and the Combined Method are shown in Fig. 4.23 as two dimensional maps binned in momentum and pseudorapidity. As with the Long Method alone, almost all bins show a ratio close to one which means that the tracking efficiency in simulation and in collision data is comparable.

#### 4.8.3. Bremsstrahlung and hadronic interactions

The tracking efficiency presented in this thesis is strictly speaking the efficiency to reconstruct a muon as a long track given a standalone track in the muon system. While it has been shown in simulation that this method reproduces the tracking efficiency with the hit-based method definition for muons with a good precision, this is not necessarily the case for other particles.

The effect of bremsstrahlung will play a role for electrons. This is the dominant source of energy loss for this type of particle and it can lead to kinks in its flight path which will reduce the quality of the track or even render it impossible to reconstruct.

The dominant effect for hadrons are hadronic interactions. As seen in Sect. 3.2.2 this often results in inelastic scatterings including the loss of the initial hadron. A simulation of  $B^0 \rightarrow J/\psi K^{*0}$  decays shows that about 11% of the tracks from kaons and about 14% of

<sup>23</sup>It is set to zero if the difference is covered by the statistical uncertainties.

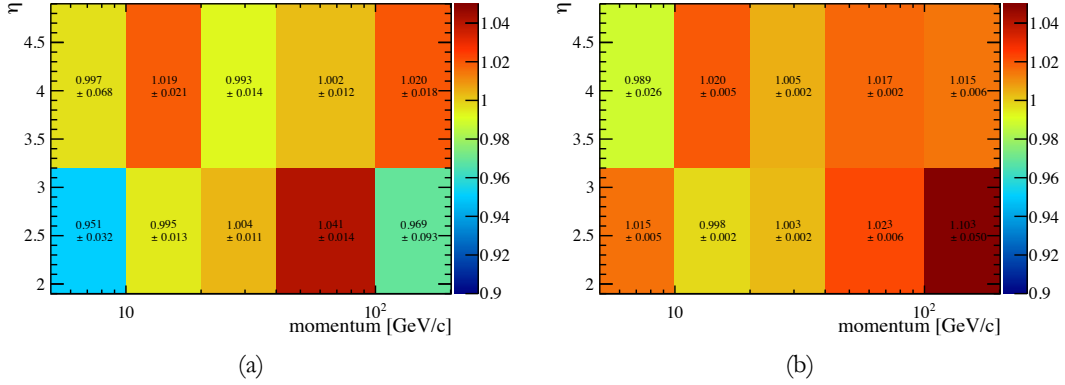


Figure 4.23.: (a) and (b) show the combined ratio between the tracking efficiency in collision data and in (reweighted) simulation determined with the Long Method and the Combined Method for 2010 and 2011 data, respectively. The additional systematic uncertainty of 0.4%, as described in Sect. 4.8.1, is not included in the quoted uncertainties.

the tracks from pions cannot be reconstructed because the particles underwent a hadronic interaction before the T stations.

Although these effects are not measured with the presented tracking efficiency determination, they are accounted for in the simulation. The corresponding tracking efficiencies can therefore be obtained from simulated samples. Additional uncertainties, however, are the total amount of relevant material in the detector, which is only known up to 10%, and the cross-sections of the hadronic interactions. To account for these effects, an uncertainty of 1.1% to 1.4% is added for the tracking efficiency obtained from simulation for hadrons.



## 5. Theoretical introduction to

$$B^0 \rightarrow K^{*0} \mu^+ \mu^-$$

### 5.1. $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ : An intriguing decay

The rare decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is a flavour changing neutral current decay<sup>1</sup> and therefore can only proceed via loop-diagrams in the Standard Model, see Chap. 1. This fact makes this decay particularly interesting as «new» physics<sup>2</sup> can enter at the same level as Standard Model physics.<sup>3</sup> Three Feynman diagrams contribute at lowest order to  $B^0 \rightarrow K^{*0} \ell \ell$ , with  $\ell$  denoting a lepton, in the Standard Model: An electromagnetic penguin diagram, exchanging a photon; a weak penguin diagram, exchanging a  $Z^0$  boson; and a box diagram, exchanging two  $W$ s. These diagrams are shown in Fig. 5.1. Furthermore, a supersymmetric diagram that could contribute to the decay is shown. The different Standard Model diagrams contributing to the decay make it particularly interesting. This is in contrast to *e.g.*  $B^0 \rightarrow K^{*0} \gamma$  which can only proceed via an electromagnetic penguin.

The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  was first observed in 2003 at BaBar [114] and since then has been further analysed by BaBar [115], Belle [116] and CDF [117].

A large amount of literature about  $B^0 \rightarrow K^{*0} \ell \ell$  and  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  exists and this chapter will only give a brief introduction to the most relevant aspects of this thesis. For more detailed information, Refs. [118] [119] [120] [121] [122] [123] [124] should be considered.

### 5.2. Decay topology

With a decay length  $c\tau = 457.2 \mu\text{m}$ , the  $B^0$  has a comparably long lifetime. Taking also the Lorentz boost into account, the  $B^0$  usually travels some millimetres from its creation to its decay when produced at a proton-proton collision at the LHC. This enables the track reconstruction to clearly separate the decay vertex of the  $B^0$  from the primary vertex. The  $K^{*0}$  however only has a lifetime of  $\tau \approx \frac{\hbar}{\Gamma} = 10^{-23}\text{s}$  and therefore decays instantly in terms of reconstruction in a  $K^+$  and a  $\pi^-$ . The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  thus manifests itself as a four-prong decay with a secondary vertex displaced from the primary.

<sup>1</sup>If not stated otherwise, the  $\mathcal{CP}$  conjugated decays are implied throughout this thesis.

<sup>2</sup>New physics refers to any process which proceeds in ways not predicted in the Standard Model.

<sup>3</sup>If new physics would enter at the tree level, it would have been seen before unless strongly suppressed couplings and very heavy particles would be present in the interaction.

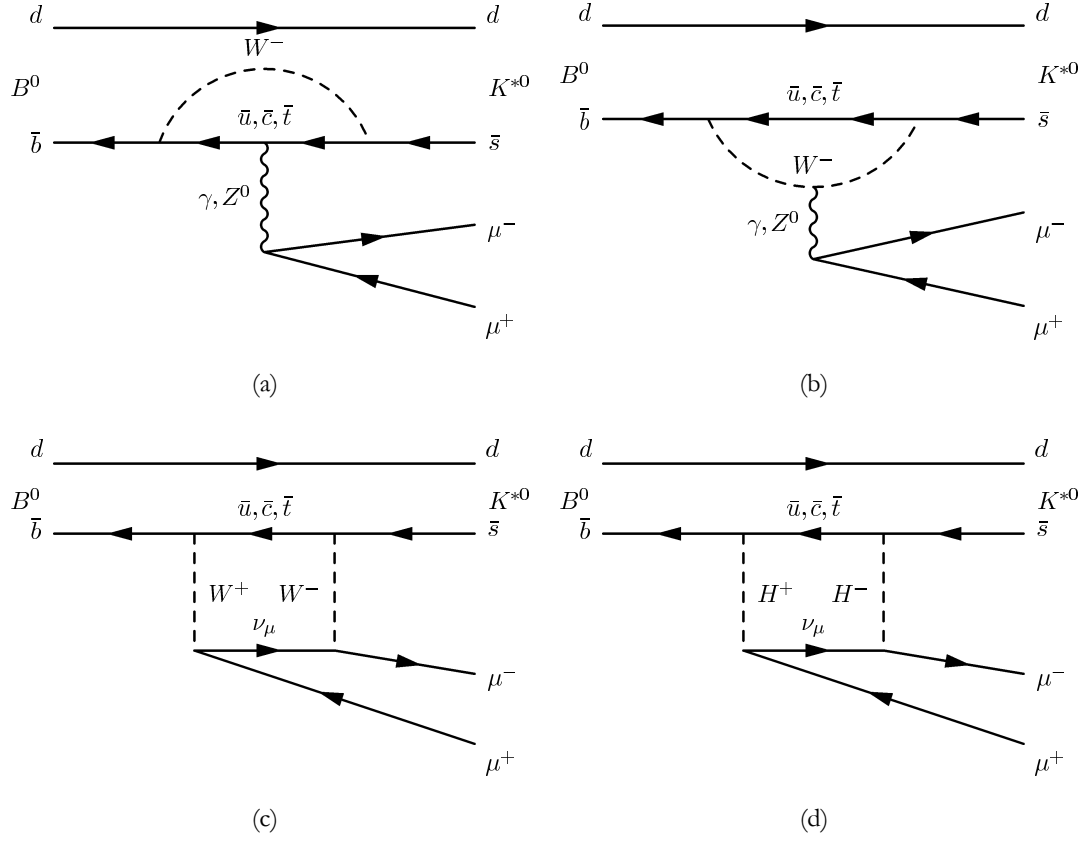


Figure 5.1.: Feynman diagrams for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . (a) and (b) depict Standard Model penguin processes, (c) shows a Standard Model box diagram while (d) shows a supersymmetric box diagram with charged Higgs exchange.

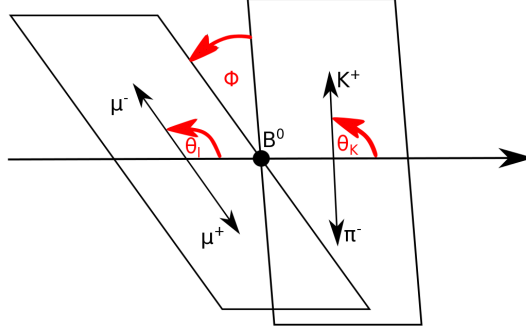


Figure 5.2.: Definition of the angles in the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ .

### 5.2.1. Definition of the angles

The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  can be fully described by four variables:  $q^2$ , the dimuon invariant mass squared; and the three angles  $\theta_K$ ,  $\theta_\ell$  and  $\phi$  where  $\theta_K$  is the angle between the  $B^0$  flight direction and the  $K^+$  in the  $K^{*0}$  rest frame,  $\theta_\ell$  is the angle between the  $B^0$  flight direction and the  $\mu^+$  in the dimuon rest frame, and  $\phi$  is the angle between the decay planes of the  $K^{*0}$  and the dimuon system.<sup>4</sup> The angles are depicted in Fig. 5.2.

The angles can also be expressed more formally: Let  $a$  be a particle and  $f$  be a rest frame. Then

$$\vec{p}_a^f \quad (5.1)$$

denotes the momentum of particle  $a$  in the rest frame of  $f$ . Furthermore define:

$$\vec{p}_{ab}^f = \vec{p}_a^f + \vec{p}_b^f \quad (5.2)$$

$$\vec{q}_{ab}^f = \vec{p}_a^f - \vec{p}_b^f \quad (5.3)$$

$$\hat{n}_{ab}^f = \frac{\vec{p}_a^f \times \vec{p}_b^f}{|\vec{p}_a^f \times \vec{p}_b^f|}, \quad (5.4)$$

with  $b$  a particle different than  $a$  and  $ab$  the common mother particle of  $a$  and  $b$ . The angles can then be expressed as:

<sup>4</sup>Note that these angle definitions are not consistent throughout literature.

$$\cos \theta_\ell = \frac{\vec{p}_{\mu^+\mu^-}^{\mu^+\mu^-} \cdot \vec{p}_{\mu^+\mu^-}^B}{\left| \vec{p}_{\mu^+\mu^-}^{\mu^+\mu^-} \right| \left| \vec{p}_{\mu^+\mu^-}^B \right|} \quad (5.5)$$

$$\cos \theta_K = \frac{\vec{p}_K^{K\pi} \cdot \vec{p}_K^B}{\left| \vec{p}_K^{K\pi} \right| \left| \vec{p}_K^B \right|} \quad (5.6)$$

$$\cos \phi = \hat{n}_{\mu^+\mu^-}^{B^0} \cdot \hat{n}_{K^+\pi^-}^{B^0} \quad \text{and} \quad (5.7)$$

$$\sin \phi = \left( \hat{n}_{\mu^+\mu^-}^{B^0} \times \hat{n}_{K^+\pi^-}^{B^0} \right) \cdot \frac{\vec{p}_{K^+\pi^-}^{B^0}}{\left| \vec{p}_{K^+\pi^-}^{B^0} \right|} \quad (5.8)$$

Note that while the definitions for  $\cos \theta_\ell$  and  $\cos \theta_K$  stay the same when exchanging the  $B^0$  with a  $\bar{B}^0$  (*i.e.* applying a  $\mathcal{CP}$  transformation), this is not true for  $\phi$ . In this angle the definitions change to:

$$\cos \phi = -\hat{n}_{\mu^+\mu^-}^{\bar{B}^0} \cdot \hat{n}_{K^-\pi^+}^{\bar{B}^0} \quad \text{and} \quad (5.9)$$

$$\sin \phi = \left( \hat{n}_{\mu^+\mu^-}^{\bar{B}^0} \times \hat{n}_{K^-\pi^+}^{\bar{B}^0} \right) \cdot \frac{\vec{p}_{K^-\pi^+}^{\bar{B}^0}}{\left| \vec{p}_{K^-\pi^+}^{\bar{B}^0} \right|} \quad (5.10)$$

**A note on the bases:** Two bases are commonly used to describe the decay of a pseudoscalar particle to two vector particles: The «helicity basis» and the «transversity basis». As the pseudoscalar particle has spin 0, the projection of the spin on the decay axis of the final state also has to be zero. In other words: Both decay products need to have the same helicity<sup>5</sup>. For (massive) vector particles, there are three possible helicity states:  $+1, 0, -1$ . For each helicity state, there is one complex amplitude: a longitudinal amplitude  $A_0$  and two transverse amplitudes  $A_{-1}, A_{+1}$ . These three amplitudes define the helicity basis. While the longitudinal amplitude is  $\mathcal{CP}$ -odd, the transverse ones are a mixture of  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd components. To remedy this flaw, the «transversity basis» can be defined. While  $A_0$  stays the same,  $A_\parallel$  and  $A_\perp$  are introduced as:

$$A_\parallel = \frac{A_{+1} + A_{-1}}{\sqrt{2}} \quad (5.11)$$

$$A_\perp = \frac{A_{+1} - A_{-1}}{\sqrt{2}} \quad (5.12)$$

$A_\parallel$  now is  $\mathcal{CP}$ -even and  $A_\perp$  is  $\mathcal{CP}$ -odd.

The angles of the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  are given in the helicity basis, the decay amplitudes however are given in the transversity basis (see next section).

---

<sup>5</sup>Helicity is defined as the normalised projection of the spin along the momentum of the particle, and, as the spin component along an axis is discrete, the helicity also takes discrete values.

### 5.2.2. Differential decay rate

The differential decay rate of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  can be expressed as a sum of products of coefficients (containing the information of the  $K^{*0}$  decay amplitudes) and angular expressions of  $\theta_\ell$ ,  $\theta_K$  and  $\phi$ . A derivation is given in Ref. [118]. When neglecting the lepton masses, the differential decay rate of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  can be written as<sup>6</sup>:

$$\begin{aligned} \frac{d^4\Gamma}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} = & \frac{9}{32\pi} \left( I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + \right. & (5.13) \\ & (I_2^s \sin^2\theta_K + I_2^c \cos^2\theta_K) \cos 2\theta_\ell + \\ & I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + \\ & I_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi + \\ & I_5 \sin 2\theta_K \sin \theta_\ell \cos \phi + \\ & I_6 \sin^2\theta_K \cos \theta_\ell + \\ & I_7 \sin 2\theta_K \sin \theta_\ell \sin \phi + \\ & I_8 \sin 2\theta_K \sin 2\theta_\ell \sin \phi + \\ & \left. I_9 \sin^2\theta_K \sin^2\theta_\ell \sin 2\phi \right) \end{aligned}$$

with the  $I_i$ s as follows:

---

<sup>6</sup>This also neglects an S-wave component coming from a non-resonant decay  $B^0 \rightarrow K^+ \pi^- \mu^+ \mu^-$  or pollution from higher  $K^*$  states like  $K_0^*(1430)$ . However, this is more of an experimental issue, as in collision data, resonant and non-resonant  $K^+ \pi^-$  cannot be disentangled.

$$\begin{aligned}
I_1^c &= (|A_{0L}|^2 + |A_{0R}|^2) \\
I_1^s &= \frac{3}{4} (|A_{\parallel L}|^2 + |A_{\parallel R}|^2 + |A_{\perp L}|^2 + |A_{\perp R}|^2) \\
I_2^c &= -(|A_{0L}|^2 + |A_{0R}|^2) \\
I_2^s &= \frac{1}{4} (|A_{\parallel L}|^2 + |A_{\parallel R}|^2 + |A_{\perp L}|^2 + |A_{\perp R}|^2) \\
I_3 &= \frac{1}{2} (|A_{\perp L}|^2 - |A_{\parallel L}|^2 + |A_{\perp R}|^2 - |A_{\parallel R}|^2) \\
I_4 &= \frac{1}{\sqrt{2}} (Re(A_{0L}A_{\parallel L}^*) + Re(A_{0R}A_{\parallel R}^*)) \\
I_5 &= \sqrt{2} (Re(A_{0L}A_{\perp L}^*) - Re(A_{0R}A_{\perp R}^*)) \\
I_6 &= 2 (Re(A_{\parallel L}A_{\perp L}^*) - Re(A_{\parallel R}A_{\perp R}^*)) \\
I_7 &= \sqrt{2} (Im(A_{0L}A_{\parallel L}^*) - Im(A_{0R}A_{\parallel R}^*)) \\
I_8 &= \frac{1}{\sqrt{2}} (Im(A_{0L}A_{\perp L}^*) + Im(A_{0R}A_{\perp R}^*)) \\
I_9 &= (Im(A_{\parallel L}A_{\perp L}^*) + Im(A_{\parallel R}A_{\perp R}^*))
\end{aligned} \tag{5.14}$$

$A_0$ ,  $A_{\parallel}$  and  $A_{\perp}$  are the decay amplitudes of the  $K^{*0}$  in the transversity basis, the indices  $L$  and  $R$  denote the chirality of the lepton current. Writing the differential decay rate as a sum of products of  $I_i$  terms and angular terms has the advantage of separating two dependencies: The  $I_i$  only depend on  $q^2$  while the rest of the expression only depends on the three angles  $\cos \theta_{\ell}$ ,  $\cos \theta_K$  and  $\phi$ .

In the literature, sometimes two terms appear for  $I_6$ :  $I_6^s$  and  $I_6^c$ . However the latter can be neglected in all cases in this thesis as it is suppressed by the small lepton mass.

The given decay rate only applies for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ , not necessarily for its  $\mathcal{CP}$  conjugated decay  $\bar{B}^0 \rightarrow \bar{K}^{*0} \mu^+ \mu^-$ .

The differential decay rate for both,  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $\bar{B}^0 \rightarrow \bar{K}^{*0} \mu^+ \mu^-$ , can be written as:

$$\frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_{\ell} d \cos \theta_K d \phi dq^2} = \frac{9}{32\pi} \sum_{i=1}^9 (I_i + \bar{I}_i) f(\cos \theta_i, \cos \theta_{\ell}, \phi). \tag{5.15}$$

In the context of this thesis, only these  $\mathcal{CP}$  averaged quantities will be discussed and measured. The  $\mathcal{CP}$  average of the  $I_k$  terms is often called  $S_k$  (for symmetric) in contrast to  $A_k$ , which corresponds to the  $\mathcal{CP}$  asymmetric ones:

$$S_i = \frac{I_i + \bar{I}_i}{\Gamma + \bar{\Gamma}} \quad (5.16)$$

$$A_i = \frac{I_i - \bar{I}_i}{\Gamma + \bar{\Gamma}} \quad (5.17)$$

As can be seen from the list of the  $I_i$  in Eq. 5.14, not all terms are independent. Four of them are connected via:

$$I_1^s = 3I_2^s \quad (5.18)$$

$$I_1^c = -I_2^c \quad (5.19)$$

### 5.2.3. Definition of related observables

Although the differential decay rate is completely described by the  $I_i$  terms, three other expressions are often used in a angular measurement of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . They are the forward-backward asymmetry,  $A_{FB}$ , the longitudinal polarisation of the  $K^{*0}$ ,  $F_L$ , and the transverse polarisation asymmetry,  $A_T^{(2)}$ . In terms of transversity amplitudes, they are defined as follows:

$$A_{FB} = \frac{3}{2} \frac{\text{Re}(A_{\parallel L} A_{\perp L}^*) - \text{Re}(A_{\parallel R} A_{\perp R}^*)}{|A_{0L}|^2 + |A_{\parallel L}|^2 + |A_{\perp L}|^2 + |A_{0R}|^2 + |A_{\parallel R}|^2 + |A_{\perp R}|^2} \quad (5.20)$$

$$F_L = \frac{|A_{0L}|^2 + |A_{0R}|^2}{|A_{0L}|^2 + |A_{\parallel L}|^2 + |A_{\perp L}|^2 + |A_{0R}|^2 + |A_{\parallel R}|^2 + |A_{\perp R}|^2} \quad (5.21)$$

$$A_T^{(2)} = \frac{|A_{\perp L}|^2 - |A_{\parallel L}|^2 + |A_{\perp R}|^2 - |A_{\parallel R}|^2}{|A_{\perp L}|^2 + |A_{\parallel L}|^2 + |A_{\perp R}|^2 + |A_{\parallel R}|^2} \quad (5.22)$$

$A_{FB}$  can be intuitively understood as the asymmetry between the forward ( $\cos \theta_\ell > 0$ ) and backward ( $\cos \theta_\ell < 0$ ) going negative muons with respect to the  $B^0$  flight direction in the dimuon rest frame. Comparing these expressions with those in Eq. 5.14 and substituting the  $S_i$  for the  $I_i$  terms, one can see that the following relations hold:

$$A_{FB} = \frac{3}{4} S_6 \quad (5.23)$$

$$F_L = S_2^c \quad (5.24)$$

$$1 - F_L = S_2^s \quad (5.25)$$

### 5.2.4. Reducing the number of observables

Equation 5.13 completely describes the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and the ultimate goal should be to measure all  $S_i$  by fitting the PDF for the differential decay rate to a dataset. However, this is not feasible with only a small dataset, as the fit would become very unstable. A straightforward method to reduce the number of observables is to integrate over two angles and only give the projection of the differential decay rate for one remaining angle. The formulae are then:

$$\frac{1}{\Gamma} \frac{d^2\Gamma}{d \cos \theta_\ell dq^2} = \frac{3}{4} F_L (1 - \cos^2 \theta_\ell) + \frac{3}{8} (1 - F_L) (1 + \cos^2 \theta_\ell) + \quad (5.26)$$

$$A_{FB} \cos \theta_\ell \quad (5.27)$$

$$\frac{1}{\Gamma} \frac{d^2\Gamma}{d \cos \theta_K dq^2} = \frac{3}{2} F_L \cos^2 \theta_K + \frac{3}{4} (1 - F_L) (1 - \cos^2 \theta_\ell) \quad (5.28)$$

$$\frac{1}{\Gamma} \frac{d^2\Gamma}{d\phi dq^2} = \frac{1}{2\pi} \left( 1 + \frac{1}{2} (1 - F_L) A_T^{(2)} \cos 2\phi + A_{Im} \sin 2\phi \right) \quad (5.29)$$

The observable  $A_{Im}$  is defined as:

$$A_{Im} = \frac{Im(A_{\parallel L} A_{\perp L}^*) + Im(A_{\parallel R} A_{\perp R}^*)}{|A_{0L}|^2 + |A_{\parallel L}|^2 + |A_{\perp L}|^2 + |A_{0R}|^2 + |A_{\parallel R}|^2 + |A_{\perp R}|^2}. \quad (5.30)$$

The «projections» for  $\cos \theta_\ell$  and  $\cos \theta_K$  were employed for the LHCb measurement in 2011 [125], based on a subset of the full 2011 data set.

A more sophisticated method to reduced the number of observables directly is to «fold» the differential decay rate and exploit symmetries in the angular expressions. An example is the transformation  $\phi \rightarrow \phi + \pi$  if  $\phi < 0$ . This folding cancels all terms with a  $\cos \phi$  and  $\sin \phi$  dependency, but leaves terms with  $\cos 2\phi$  and  $\sin 2\phi$  unaffected. This transformation was used for the measurement of  $A_{FB}$ ,  $F_L$ ,  $S_3$ ,  $S_9$  which is described in Ref. [126]<sup>7</sup>. The calculation of the transformation is given in Eq. 5.43 and the following formulae.

In this thesis, only the (remaining) observables  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  are extracted. The following transformations are applied to obtain decay rates that only contain  $F_L$ ,  $A_T^{(2)}$  and the corresponding  $S_i$  observables as parameters.<sup>8</sup>

**Transformations for  $S_4$**  The following transformations are applied to obtain a formula with only  $S_4$  among the  $S_i$  observables:

<sup>7</sup>Additionally the observables  $A_T^{(2)}$ ,  $A_T^{Re}$  and  $A_9$  are measured in Ref. [126], where also their definitions are given.

<sup>8</sup>It is in principle possible to fold these PDFs even more to get rid of  $A_T^{(2)}$ . However this comes with a loss of sensitivity to the  $S_i$  observables which is not desired.



$$\phi \rightarrow -\phi \quad \text{if } \phi < 0 \quad (5.31)$$

$$\phi \rightarrow \pi - \phi \quad \text{if } \theta_\ell > \frac{\pi}{2} \quad (5.32)$$

$$\theta_\ell \rightarrow \pi - \theta_\ell \quad \text{if } \theta_\ell > \frac{\pi}{2}. \quad (5.33)$$

This leads to the following differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ & + \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & \left. + S_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \right) \end{aligned}$$

**Transformations for  $S_5$**  The following transformations are applied to obtain a formula with only  $S_5$  among the  $S_i$  observables:

$$\phi \rightarrow -\phi \quad \text{if } \phi < 0 \quad (5.34)$$

$$\theta_\ell \rightarrow \pi - \theta_\ell \quad \text{if } \theta_\ell > \frac{\pi}{2} \quad (5.35)$$

This leads to the following differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ & + \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & \left. + S_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \right) \end{aligned}$$

**Transformations for  $S_7$**  The following transformations are applied to obtain a formula with only  $S_7$  among the  $S_i$  observables:

$$\phi \rightarrow \pi - \phi \quad \text{if } \phi > \frac{\pi}{2} \quad (5.36)$$

$$\phi \rightarrow -\pi - \phi \quad \text{if } \phi < -\frac{\pi}{2} \quad (5.37)$$

$$\theta_\ell \rightarrow \pi - \theta_\ell \quad \text{if } \theta_\ell > \frac{\pi}{2} \quad (5.38)$$

This leads to the following differential decay rate::

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} &= \frac{9}{8\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ &+ \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ &+ \frac{1}{2} (1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ &\left. + S_7 \sin 2\theta_K \sin \theta_\ell \sin \phi \right) \end{aligned}$$

**Transformations for  $S_8$**  The following transformations are applied to obtain a formula with only  $S_8$  among the  $S_i$  observables:

$$\phi \rightarrow \pi - \phi \quad \text{if } \phi > \frac{\pi}{2} \quad (5.39)$$

$$\phi \rightarrow -\pi - \phi \quad \text{if } \phi < -\frac{\pi}{2} \quad (5.40)$$

$$\theta_\ell \rightarrow \pi - \theta_\ell \quad \text{if } \theta_\ell > \frac{\pi}{2} \quad (5.41)$$

$$\theta_K \rightarrow \pi - \theta_K \quad \text{if } \theta_\ell > \frac{\pi}{2} \quad (5.42)$$

This leads to the following differential decay rate::

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} &= \frac{9}{8\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ &+ \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ &+ \frac{1}{2} (1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ &\left. + S_8 \sin 2\theta_K \sin 2\theta_\ell \sin \phi \right) \end{aligned}$$

All four PDFs can be fitted to data and will yield the value for  $S_i$  and additionally  $F_L$  and  $A_T^{(2)}$ . Although the latter observables are also of interest in a general context, they will be

rather treated as «nuisance parameters», *i.e.* parameters which appear in the PDF but whose value is a by-product rather than a result. Their value was already measured in a different analysis of LHCb [126] with the same 2011 data set, which is briefly described in Chap. 7. The full derivation of the formulae for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  is given in Appendix D.

**Transformations to obtain  $F_L$ ,  $A_{FB}$ ,  $S_3$  and  $S_9$**  The following transformations are applied to obtain a formula with  $F_L$ ,  $A_{FB}$ ,  $S_3$  and  $S_9$ :

$$\phi \rightarrow \phi + \pi \quad \text{if } \phi < 0 \quad (5.43)$$

This leads to the following differential decay rate::

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{16\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ & + \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + S_3 \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & + \frac{4}{3} A_{FB} \sin^2 \theta_K \cos \theta_\ell \\ & \left. + S_9 \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \right) \end{aligned}$$

The additional observables and how they can be obtained is described in Ref. [126].

### 5.2.5. Extracting the $S_i$ observables using a counting experiment

The  $S_i$  observables can in principle also be extracted using a simple binned counting method or by fitting the differential branching fraction: as a function of  $q^2$  for two classes of events («positive» and «negative») and then constructing an asymmetry. For example,  $S_5$  can be constructed as:

$$\begin{aligned} S_5 = & \frac{4}{3} \left( \int_{-\pi/2}^{\pi/2} - \left( \int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi} \right) d\phi \right) \left( \int_0^\pi d\theta_\ell \right) \times \\ & \left( \int_0^{\pi/2} - \int_{\pi/2}^\pi d\theta_K \right) \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} \end{aligned} \quad (5.44)$$

and then be split into two classes:

$$\begin{aligned}
S_5^+ &= \frac{4}{3} \left( \int_{-\pi/2}^{\pi/2} d\phi \right) \left( \int_0^\pi d\theta_\ell \right) \left( \int_0^{\pi/2} d\theta_K \right) \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} \\
&+ \frac{4}{3} \left( \left( \int_{-\pi}^{-\pi/2} + \int_{\pi/2}^\pi \right) d\phi \right) \left( \int_0^\pi d\theta_\ell \right) \left( \int_{\pi/2}^\pi d\theta_K \right) \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2}
\end{aligned} \tag{5.45}$$

$$\begin{aligned}
S_5^- &= \frac{4}{3} \left( \int_{-\pi/2}^{\pi/2} d\phi \right) \left( \int_0^\pi d\theta_\ell \right) \left( \int_{\pi/2}^\pi d\theta_K \right) \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} \\
&+ \frac{4}{3} \left( \left( \int_{-\pi}^{-\pi/2} + \int_{\pi/2}^\pi \right) d\phi \right) \left( \int_0^\pi d\theta_\ell \right) \left( \int_0^{\pi/2} d\theta_K \right) \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2}
\end{aligned} \tag{5.46}$$

Both,  $S_5^+$  and  $S_5^-$ , can be determined by fitting the corresponding distributions in  $q^2$  similarly to what is done for the extraction of the zero-crossing point, see Chap. 8. However, while this procedure is easier to accomplish than fitting the formulae given in 5.36, it has the disadvantage of a reduced sensitivity and a reduced range in  $q^2$ , as the full kinematical region in  $q^2$  cannot be accessed. More details on the method for extracting the values of  $S_i$  without using an angular fit can be found in Ref. [127].

While binned counting in principle is possible for the full  $q^2$  range, the reduced sensitivity is even more of a concern. A comparison of the sensitivity between binned counting and an angular fit is shown as an example for  $S_5$  in Fig. 5.3. The situation is similar for all  $S_i$  observables.

### 5.2.6. Introducing an S-wave system

The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  does not have an S-wave component, as the  $K^+ \pi^-$  system is bound in the  $K^{*0}$  system. However, experimentally a  $K^+ \pi^-$  system that happens to have been in the invariant mass region of the  $K^{*0}$  cannot be distinguished from the bound system<sup>9</sup>. The total differential decay rate can be written as:

$$\frac{d^4\Gamma}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} \propto (1 - F_S) PDF_{K^{*0}} + PDF_S \tag{5.47}$$

where  $F_S = \frac{\Gamma_S}{\Gamma_{tot}}$  is the fraction of the S-wave decay rate with respect to the total decay rate,  $PDF_{K^{*0}}$  is the distribution of Eq. 5.13 and  $PDF_S$  is the distribution for the S-wave component and its interference with the P-wave from the  $K^{*0}$ . It is defined as:

---

<sup>9</sup>This  $K^+ \pi^-$  may come from non-resonant decays or be a contribution from a  $K_0^{*0}(1430)$  decay.

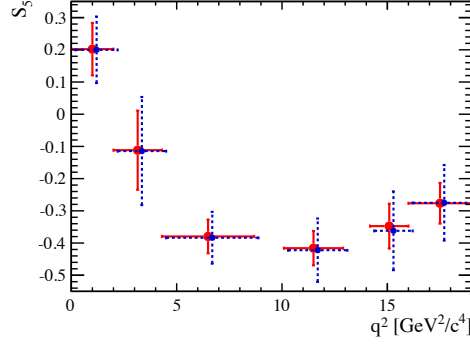


Figure 5.3.: Comparison of the sensitivity to  $S_5$ . In blue dashed are the  $S_5$  measured with a binned counting method, in red solid are the values obtained with a fit to the differential branching fraction. The result was obtained performing toy-experiments, using the Standard Model predictions as an input and the number of observed events per bin taken from Ref. [126]. The gaps around  $9 \text{ GeV}^2/c^4$  and  $13.5 \text{ GeV}^2/c^4$  stem from the vetoes on  $B^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow \psi(2S)K^{*0}$ .

$$\begin{aligned}
 PDF_S = & \frac{2}{3}F_S \sin^2 \theta_\ell + \frac{4}{3}A_S \sin^2 \theta_\ell \cos \theta_K + \\
 & A_S^{(4)} \sin \theta_K \sin 2\theta_\ell \cos \phi + \\
 & A_S^{(5)} \sin \theta_K \sin \theta_\ell \cos \phi + \\
 & A_S^{(7)} \sin \theta_K \sin \theta_\ell \sin \phi + \\
 & A_S^{(8)} \sin \theta_K \sin 2\theta_\ell \sin \phi.
 \end{aligned} \tag{5.48}$$

The  $A_S$  and  $A_S^{(i)}$  expressions are interference terms between the S-wave and the P-wave state and are defined in Appendix E. These terms have upper and lower boundaries:

$$|A_S| \leq \sqrt{3}\sqrt{F_S(1-F_S)F_L} \tag{5.49}$$

$$|A_S^{(4,8)}| \leq \sqrt{\frac{2}{3}}\sqrt{F_S(1-F_S)(1-F_L)} \tag{5.50}$$

$$|A_S^{(5,7)}| \leq \sqrt{\frac{8}{3}}\sqrt{F_S(1-F_S)(1-F_L)} \tag{5.51}$$

The formulae for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$ , including the S-wave system and applying the folding, are then:

**Formula for  $S_4$ :**

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{2}{3} F_S \sin^2 \theta_\ell + \frac{4}{3} A_S \cos \theta_K \sin^2 \theta_\ell \right. \\ & + \frac{3}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K + (1 - F_S) F_L \cos^2 \theta_K \\ & + \frac{1}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell \\ & - (1 - F_S) F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_S)(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & + (1 - F_S) S_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \\ & \left. + A_S^{(4)} \sin \theta_K \sin 2\theta_\ell \cos \phi \right) \end{aligned} \quad (5.52)$$

**Formula for  $S_5$ :**

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{2}{3} F_S \sin^2 \theta_\ell + \frac{4}{3} A_S \cos \theta_K \sin^2 \theta_\ell \right. \\ & + \frac{3}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K + (1 - F_S) F_L \cos^2 \theta_K \\ & + \frac{1}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell \\ & - (1 - F_S) F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_S)(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & + (1 - F_S) S_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \\ & \left. + A_S^{(5)} \sin \theta_K \sin \theta_\ell \cos \phi \right) \end{aligned} \quad (5.53)$$

**Formula for  $S_7$ :**

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{2}{3} F_S \sin^2 \theta_\ell + \frac{4}{3} A_S \cos \theta_K \sin^2 \theta_\ell \right. \\ & + \frac{3}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K + (1 - F_S) F_L \cos^2 \theta_K \\ & + \frac{1}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell \\ & - (1 - F_S) F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_S)(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & + (1 - F_S) S_7 \sin 2\theta_K \sin \theta_\ell \sin \phi \\ & \left. + A_S^{(7)} \sin \theta_K \sin \theta_\ell \sin \phi \right) \end{aligned} \quad (5.54)$$

### Formula for $S_8$ :

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{2}{3} F_S \sin^2 \theta_\ell \right. \\ & + \frac{3}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K + (1 - F_S) F_L \cos^2 \theta_K \\ & + \frac{1}{4} (1 - F_S)(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell \\ & - (1 - F_S) F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2} (1 - F_S)(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & + (1 - F_S) S_8 \sin 2\theta_K \sin 2\theta_\ell \sin \phi \\ & \left. + A_S^{(7)} \sin \theta_K \sin 2\theta_\ell \sin \phi \right) \end{aligned} \quad (5.55)$$

The experimental use of these formulae is explained in Sect. 9.4.

## 5.3. Theoretical uncertainties and clean observables

The process of interest in  $B^0 \rightarrow K^{*0} \ell \ell$  is the transition  $b \rightarrow s \ell^+ \ell^-$ . However, as quarks are confined within a meson, this transition cannot be observed directly and the exclusive decay  $B^0 \rightarrow K^{*0} \ell \ell$  has to be measured. A complication arises as bound objects like mesons involve non-perturbative QCD effects, which are not calculable exactly. The transition  $B^0 \rightarrow K^{*0}$  can be parametrised using so-called «form factors». Some form factors are observables while others depend on the choice of the renormalisation scale  $\mu$ . More details are given in Ref. [118]. The uncertainties on the form factors contribute the largest part of the theoretical uncertainty.

It was pointed out in Refs. [123] [124] that a maximum of six form factor independent observables<sup>10</sup> can be formed. The decay is then completely described by the set  $\left\{ \frac{d\Gamma}{dq^2}, F_L, P_1, \dots, P_6 \right\}$ : while  $\frac{d\Gamma}{dq^2}$  and  $F_L$  suffer from form factor uncertainties, the  $P_i$  do not, which allows to predict them more precisely. These  $P_i$  observables are related to the  $S_i$  mentioned before. However, while  $P_1, P_2$ <sup>11</sup> and  $P_3$  can be measured experimentally, it is advantageous to slightly redefine  $P_4, P_5$  and  $P_6$  to  $P'_4, P'_5$  and  $P'_6$ . The primed quantities are more easily measurable as they depend less on averaging different observables over the bin size. Furthermore they are as clean as the non-primed ones. To make the comparison

<sup>10</sup>Strictly speaking, these observables are only independent of form factor uncertainties in leading order and at low  $q^2$  although for some of them this is also true for high  $q^2$ .

<sup>11</sup>Although bearing the same name, the observable  $P_2$  is not linked in any way to the Italian (pseudo-)Masonic lodge known as «Propaganda Due».

between experimental and theoretical results the most discriminating, the idea is to measure all  $P_i^{(t)}$  observables.

A different approach to access information which does not suffer from large theoretical uncertainties is to measure zero-crossings of  $S_i$  observables. Although the  $S_i$  observables are affected by form-factor uncertainties, they cancel at the point of the zero-crossing in leading-order. This is for example exploited in the measurement of the zero-crossing point of  $A_{FB}$ . However, while the cancellation for  $A_{FB}$  only occurs in one point, the cancellation for the  $P_i^{(t)}$  observables is valid in the full  $q^2$  range.

### 5.3.1. Definition of the $P'_i$ observables

As explained in the previous section it is advantageous to measure the  $P'_i$  observables instead of the  $S_i$  ones. They relate to the  $S_i$  as follows<sup>12</sup>:

$$P'_4 = \frac{S_4}{\sqrt{F_L(1 - F_L)}} \quad (5.56)$$

$$P'_5 = \frac{S_5}{\sqrt{F_L(1 - F_L)}} \quad (5.57)$$

$$P'_6 = \frac{S_7}{\sqrt{F_L(1 - F_L)}} \quad (5.58)$$

$$P'_8 = \frac{S_8}{\sqrt{F_L(1 - F_L)}} \quad (5.59)$$

The three observables  $P_1, P_2, P_3$  on the other hand are related to  $S_3, S_6, S_9$  and  $F_L$ . While the observables  $P_1, P_2, P_3, P'_4, P'_5$  and  $P'_6$  are independent observables,  $P'_8$  is not<sup>13</sup>, as it can be constructed as a combination of the other  $P_i^{(t)}$  observables,  $F_L$  and  $\frac{d\Gamma}{dq^2}$ . However its measurement is still important to check the consistency with the other observables — additionally it could show signs of new physics for the high- $q^2$  region [128].

## 5.4. Theoretical predictions

All the observables discussed in this chapter have predictions in the Standard Model and some models beyond the Standard Model. Some of them can be found in Refs. [118] [124]. However, due to the presence of the  $B^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow \psi(2S)K^{*0}$  resonance, the region in  $q^2$  between about  $9.0 \text{ GeV}^2/c^4$  and  $14.0 \text{ GeV}^2/c^4$  is difficult to predict and only

<sup>12</sup>The definition here is slightly different from the one in [124]. The definition in this thesis was chosen to render the definitions more consistent among the observables. However, the difference only consists of minus signs and factors of 2.

<sup>13</sup>To add to the confusion,  $P'_8$  is sometimes also called  $Q$ . Or  $H_T^{(4)}$ . Bazinga!



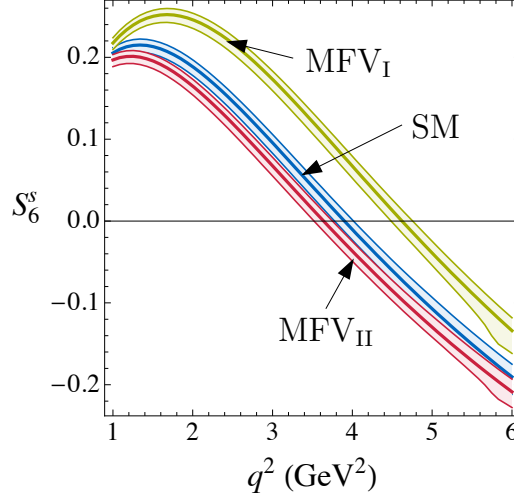


Figure 5.4.: Prediction for  $S_6^s$  ( $= \frac{4}{3}A_{FB}$ ) in the Standard Model (blue) and two Minimal Flavour Violating Scenarios (green and red), see the text. Figure taken from Ref. [118].

the regions for low  $q^2$  and high  $q^2$  are normally given. Furthermore, the region below  $q^2$  of 1  $\text{GeV}^2/c^4$  is often omitted due to the presence of the «photon-pole», which leads to a divergence in the differential branching fraction<sup>14</sup>, and due to the presence of other resonances like  $\rho(770)$ . The theoretically cleanest region is normally considered between 1.0 - 6.0  $\text{GeV}^2/c^4$ , and most observables are predicted in this interval.

As an example, Fig. 5.4 shows different predictions for  $A_{FB}$  in the region of  $q^2$  between 1.0 - 6.0  $\text{GeV}^2/c^4$  for the Standard Model and two Minimal Flavour Violating scenarios<sup>15</sup>.

## 5.5. Measurements of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ by CDF and the $B$ -factories

Before the start of LHCb, the B-factory experiments Belle and BaBar, and additionally CDF, had analysed  $B^0 \rightarrow K^{*0} \ell \ell$ , see Refs. [115] [116] [117]. Belle and BaBar could analyse  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  as well as  $B^0 \rightarrow K^{*0} e^+ e^-$ , while CDF analysed  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ . However, these three experiments only had about 230, 60 and 164 events in their dataset respectively. No fit in more than one angle was possible with this

<sup>14</sup>For  $q^2 \rightarrow 0$  the cross section becomes infinite. However, this is only true when neglecting the lepton masses.

<sup>15</sup>Minimal Flavour Violating assumes that flavour violation in models beyond the Standard Model are linked to Yukawa couplings as is the case in the Standard Model. More information can be found in Ref. [7].

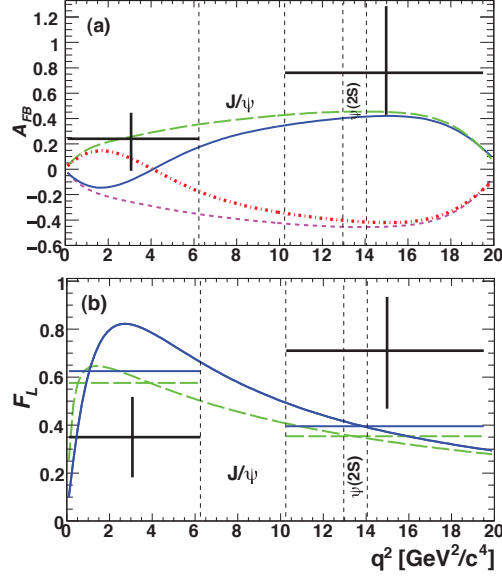


Figure 5.5.:  $A_{FB}$  and  $F_L$  as measured by the BaBar experiment, see Ref. [115]. The data include  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B^0 \rightarrow K^{*0} e^+ e^-$ . The crosses are the fit results from collision data, the blue solid line is the Standard Model prediction, the green long dashed line is the SM with  $\mathcal{C}_7^{(\text{eff})} \rightarrow -\mathcal{C}_7^{(\text{eff})}$ , the magenta short dashed line is the SM with  $\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})} \rightarrow -\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})}$  and the red dash-dotted line is  $\mathcal{C}_7^{(\text{eff})} \rightarrow -\mathcal{C}_7^{(\text{eff})}$ ,  $\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})} \rightarrow -\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})}$ . The  $F_L$  curves with  $\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})} \rightarrow -\mathcal{C}_9^{(\text{eff})} \mathcal{C}_{10}^{(\text{eff})}$  are almost identical to the two curves shown.

small number of candidates, so all three experiments used projections in one of the three angles to extract the physics observables. The results for all three experiments are shown in Figs. 5.5, 5.6 and 5.7.

The BaBar and Belle results seemed to hint a deviation from the Standard Model in the  $A_{FB}$  distribution at low  $q^2$ , while the latest data from CDF could neither confirm nor refute this finding. LHCb was the first experiment to rule it out with a large significance (see Chap. 7).

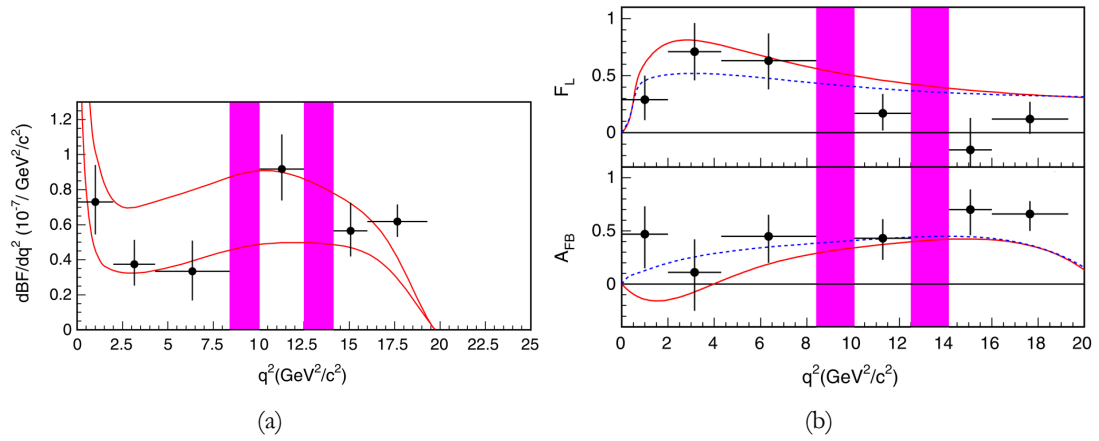


Figure 5.6.: Results of the Belle experiment for the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B^0 \rightarrow K^{*0} e^+ e^-$ , see Ref. [116]. (a) shows the differential branching fraction as function of  $q^2$ . The two red lines are the SM prediction for minimum and maximum allowed form factors. (b) shows  $F_L$  and  $A_{FB}$ : the red solid line is the SM prediction, the blue dashed line the SM prediction, but with  $C_7 = -C_7^{SM}$ . The two bands in magenta show the resonant regions of the  $J/\psi$  and  $\psi(2S)$  which were excluded.

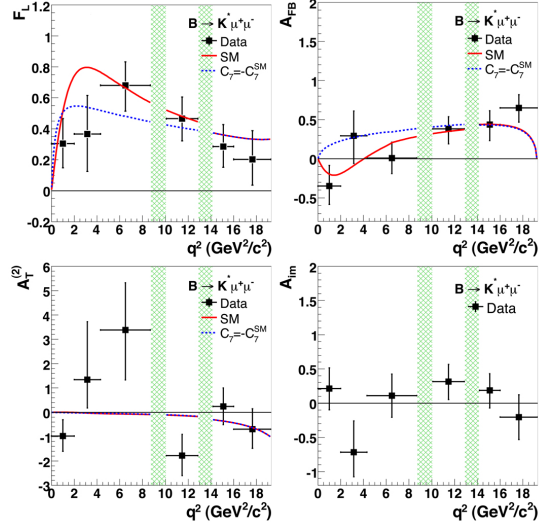


Figure 5.7.:  $F_L$ ,  $A_{FB}$ ,  $A_T^{(2)}$  and  $A_{Im}$  as measured by the CDF experiment, see Ref. [117]. The data include  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ . The points are the fit results from collision data. The SM expectation is shown as the red solid curve, the blue dashed curve is the case  $\mathcal{C}_7 = -\mathcal{C}_7^{SM}$ . The green bands show the resonant regions of the  $J/\psi$  and  $\psi(2S)$  which were excluded.

## 6. Selecting $B^0 \rightarrow K^{*0} \mu^+ \mu^-$

With a branching fraction of  $\mathcal{B} = (1.05_{-0.13}^{+0.16}) \cdot 10^{-6}$  [129] the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is rare and requires a sophisticated selection procedure to reduce the amount of background to a minimum. Furthermore, several stages of the analysis rely on simulated data: Differences in the simulation and the collision data have to be understood and corrected. This chapter presents the steps from the raw data to the final sample («nTuple») which was used for all subsequent analyses presented in this thesis and in Ref. [126]. The amount of data in this sample corresponds to roughly  $1 \text{ fb}^{-1}$  and is the full data sample collected in 2011.

The work presented in this chapter is the result of a collaborative effort and was not done solely by the author. However, all steps are mentioned for completeness.

### 6.1. Selection

The selection for the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  proceeds in several steps: The decay has to fire one of several trigger lines in all three trigger stages of LHCb and be selected in the central offline selection (stripping). Furthermore, a loose preselection and finally a multivariate selection was performed.

#### 6.1.1. Trigger selection

At the first trigger stage, L0, the event has to be triggered by a single muon from one of the muons in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . In HLT1, the event either has to be triggered by the single track trigger [74] or the single muon trigger [73]. It is demanded that the trigger in question was always fired by one of the muons from  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . In HLT2, more trigger lines were used: They are either topological triggers [75] or muon triggers [73]. As in HLT1, the trigger had to be fired by one or several of the final state particles of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ .

#### 6.1.2. Central offline selection («stripping selection»)

Before the data can be analysed by the user, a central offline selection is performed in order to loosely select events of interest for different analyses. For  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ , a cut-based selection was implemented with cuts on the  $K^+ \pi^- \mu^+ \mu^-$  system, the  $K^+ \pi^-$  system, the dimuon system and on all individual tracks. Furthermore, a loose criterion to identify the particles as muons called «IsMuonLoose» was put in the selection. For particles with a momentum between 3 and 6 GeV/c, it requires at least two hits in M2, M3, M4 and M5

compatible with the track; for particles with a momentum above 6 GeV/ $c$ , it requires at least three hits in all muon stations (except M1) compatible with the track. The cuts for the stripping selection are summarised in Table 6.1.

Table 6.1.: Selection cuts for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  in the stripping. The particle names stand for the systems of their corresponding daughter particle, *e.g.* the cuts under  $B^0$  are applied on the  $K^+ \pi^- \mu^+ \mu^-$  system.

Particle	Variable	Value
$B^0$	mass	$> 4580 \text{ MeV}/c$ and $< 5780 \text{ MeV}/c$
	DIRA <sup>1</sup>	$> 0.9999$
	$\chi^2_{\text{vtx}}/\text{ndof}$	$< 6$
	$\chi^2_{\text{IP}}$	$< 16$
	$\chi^2_{\text{FD}}$ <sup>2</sup>	$> 121$
$K^{*0}$	mass	$> 600 \text{ MeV}/c$ and $< 2000 \text{ MeV}/c$
	$\chi^2_{\text{vtx}}/\text{ndof}$	$< 12$
	$\chi^2_{\text{FD}}$	$> 9$
$\mu^+ \mu^-$	$\chi^2_{\text{vtx}}/\text{ndof}$	$< 12$
	$\chi^2_{\text{FD}}$	$> 9$
Track	$\chi^2_{\text{vtx}}/\text{ndof}$	$< 5$
	$\chi^2_{\text{IP}}$	$> 9$
	$p_T$	$> 250 \text{ MeV}/c$
$\mu$	isMuonLoose	true

### 6.1.3. Preselection

The preselection is applied to veto pathological events. It mainly contains fiducial cuts *e.g.* requiring the event to come from the primary vertex or the hadrons to have information from the RICH detectors. The cuts are summarised in Table 6.2.

<sup>1</sup>DIRA stands for «direction angle» and is the angle between the sum of the reconstructed momenta of all daughter particles of the  $B^0$  and the line joining the primary vertex and the decay vertex of the  $B^0$ .

<sup>2</sup>FD stands for flight distance. However the variable  $\chi^2_{\text{FD}}$  is not related to the flight distance. It is the «vertex separation  $\chi^2$ » and is a measure of the separation of two vertices. Technically it is the  $\chi^2$  value of the common vertex of all tracks in question minus the sum of the  $\chi^2$  for two distinct vertices. The name  $\chi^2_{\text{FD}}$  is kept for consistency with LHCb conventions.

Table 6.2.: Selection cuts for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  in the preselection.

Particle	Variable	Value
$\mu$	isMuon	true
$K^+$	hasRICH <sup>3</sup>	true
	$DLL_{K\pi}$ <sup>4</sup>	$> -5$
$\pi^-$	hasRICH	true
	$DLL_{K\pi}$	$< 25$
Track Pairs	$ \theta_{\text{track } 1} - \theta_{\text{track } 2} $	$> 1 \text{ mrad}$
Track	$\theta$	$< 400 \text{ mrad}$
	KL Distance <sup>5</sup>	$> 5000$
PV	$ x_{\text{PV}} - x_{\text{beam axis}} $	$< 5 \text{ mm}$
	$ y_{\text{PV}} - y_{\text{beam axis}} $	$< 5 \text{ mm}$
	$ z_{\text{PV}} - \bar{z}_{\text{interaction position}} $	$< 5 \text{ mm}$

#### 6.1.4. Multivariate selection

The final selection of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  was performed with a Boosted Decision Tree (BDT), using the software package TMVA [131]. A BDT consist of a tree-like structure of pass-fail criteria that are made insensitive to fluctuations using a technique called «boosting», which uses a weighted average of many decision trees. The BDT for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  uses the following variables for the discrimination between signal and background.

- For the  $B^0$ : The direction angle, the flight distance, the  $\chi_{\text{IP}}^2$  with respect to the primary vertex, the transverse momentum  $p_T$  and the  $\chi_{\text{vtx}}^2$ .
- For the intermediate objects ( $K^{*0}$ , dimuon-object): The flight distance with respect to the primary vertex, the  $\chi_{\text{IP}}^2$  with respect to the primary vertex, the transverse momentum and the  $\chi_{\text{vtx}}^2$ .
- For the final state particles: The  $\chi_{\text{IP}}^2$  with respect to the primary vertex,  $DLL_{K\pi}$  and  $DLL_{\mu\pi}$ .

<sup>3</sup>hasRICH is a boolean variable that equals true if the particle has information from the RICH detectors assigned to it.

<sup>4</sup>The  $DLL_{K\pi}$  (and  $DLL_{\mu\pi}$ ) variable will be explained in Sect. 6.2.2.

<sup>5</sup>KL Distance stands for «Kullback-Leibler distance», a measure which can be used to reject «clones», *i.e.* two tracks which are likely to be formed out of hits left by a single particle. For details see Ref. [130].

The BDT was trained for the signal using  $B^0 \rightarrow J/\psi K^{*0}$ , where  $J/\psi \rightarrow \mu^+ \mu^-$ , from collision data in 2010. To train the background, events from the  $B^0$  mass sidebands<sup>6</sup> in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  from 2010 collision data were used. These events were not included in the final analysis. Half of the 2010 sample, corresponding to  $19 \text{ pb}^{-1}$ , was used for training, while the other half was used to test the performance of the BDT. To determine the optimal working point of the BDT, the metric  $M = \frac{s}{\sqrt{s+b}}$  was used, where  $s$  is the number of signal events in the  $B^0$  mass window and  $b$  the number of background events in the  $B^0$  mass window. The range of the BDT output is between -1 and 1, the optimal cut was found to be at 0.1. It was checked that a variation around this value does not lead to a significant improvement of  $M$  when applying the BDT to the full 2011 data set. The total efficiency for selecting  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  events after the BDT is 0.93%. This value is calculated with respect to  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  decays produced in  $4\pi$  and includes efficiencies from the geometrical acceptance, trigger, reconstruction, stripping, the pre-selection and the BDT selection. The corresponding efficiency for selecting  $B^0 \rightarrow J/\psi K^{*0}$  events is 1.04%. The difference stems from the muon- $p_T$  dependence of the efficiency of the trigger and the lower efficiency for reconstructing the dimuon pair at low  $q^2$  values with respect to the  $q^2$  value of the  $J/\psi$ .

### 6.1.5. Charmonium Vetoes

The dimuon spectrum of the selected  $K^+ \pi^- \mu^+ \mu^-$  candidates shows two large excesses caused by the resonant decays of  $J/\psi \rightarrow \mu^+ \mu^-$  and  $\psi(2S) \rightarrow \mu^+ \mu^-$  from the decays  $B^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow \psi(2S) K^{*0}$ , respectively. Both decays are dominated by tree-level diagrams. As the resonant and non-resonant components cannot be disentangled in these regions, they are removed by cuts in the dimuon spectrum. A veto is applied between  $2946 \text{ MeV}/c^2$  and  $3176 \text{ MeV}/c^2$  in the dimuon mass to remove the  $J/\psi$  and between  $3568 \text{ MeV}/c^2$  and  $3766 \text{ MeV}/c^2$  in the dimuon mass to remove the  $\psi(2S)$ . The size of these vetoes are determined by the radiative tails of the respective decays. The lower bounds of the vetoes are extended to  $2796 \text{ MeV}/c^2$  and  $3436 \text{ MeV}/c^2$  for  $J/\psi \rightarrow \mu^+ \mu^-$  and  $\psi(2S) \rightarrow \mu^+ \mu^-$ , respectively, if  $m(K^+ \pi^- \mu^+ \mu^-) < 5230 \text{ MeV}/c^2$ . The upper bounds of these vetoes are extended to  $3216 \text{ MeV}/c^2$  and  $3806 \text{ MeV}/c^2$ , respectively, if  $5330 \text{ MeV}/c^2 < m(K^+ \pi^- \mu^+ \mu^-) < 5450 \text{ MeV}/c^2$ . The first extension is needed due to the correlation of low reconstructed  $B^0$  mass and reconstructing the dimuon mass at the low end of the radiative tail. The second extension on the other hand is necessary to take misreconstruction effects on the upper edge of the reconstructed  $B^0$  mass peak into account. The effect of the two extensions of the vetoes causes the introduction of a weight for the maximum likelihood fit for the angular analysis. The reconstructed  $B^0$  candidate

<sup>6</sup>The mass sidebands for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  are defined as follows: The lower sideband spans from  $5150 \text{ MeV}/c^2$  to  $5220 \text{ MeV}/c^2$ , the upper sideband from  $5350 \text{ MeV}/c^2$  to  $5750 \text{ MeV}/c^2$  and the signal region from  $5220 \text{ MeV}/c^2$  to  $5320 \text{ MeV}/c^2$ . This corresponds to the situation in Fig. 6.1, where the yellow line marks the lower bound of the lower mass sideband.



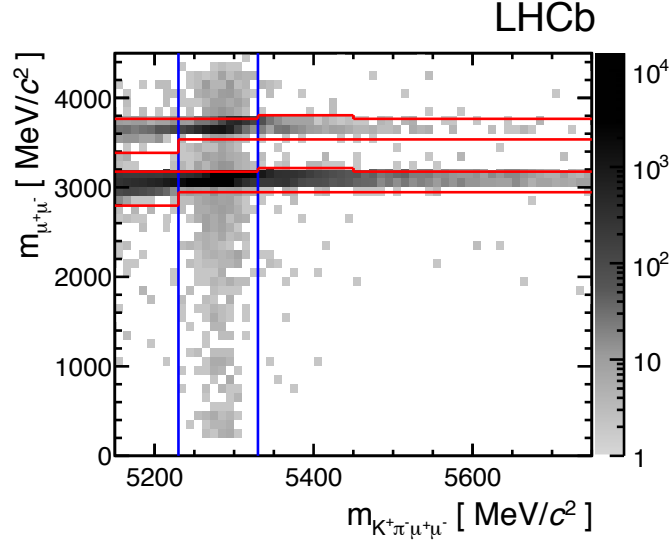


Figure 6.1.: Mass vetoes for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  to reject  $B^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow \psi(2S) K^{*0}$  as a function of the  $\mu^+ \mu^-$  and  $K^+ \pi^- \mu^+ \mu^-$  mass. In red are the vetoes in the dimuon mass spectrum and in blue is the signal region.

mass is plotted against the dimuon mass in Fig. 6.1 including all the vetoes.

### 6.1.6. Peaking backgrounds

Combinatorial background has a falling distribution as a function of the mass and can be estimated from collision data directly. Specific peaking backgrounds have to be considered separately. They consist for the largest part of exclusive decays where one or more particles are misidentified and therefore end up in the mass region of the  $B^0$ . The strategy to veto these events is the same for all of them: The event is reconstructed with different mass hypotheses for the particles, corresponding to the mass hypotheses for the peaking background decays in question. If the combination under the new hypotheses then peaks in a given region, the event is rejected.

An example is the decay  $B^0 \rightarrow J/\psi K^{*0}$ . It can be a source of peaking background, if the pion from  $K^{*0} \rightarrow K^+ \pi^-$  is misreconstructed as a muon and the  $\mu^-$  from  $J/\psi \rightarrow \mu^+ \mu^-$  is misreconstructed as a pion. The resulting invariant mass of all four final state particles then peaks in the  $B^0$  mass region. This background can be excluded by assigning the particle, reconstructed as a pion, a muon mass hypothesis and assigning the particle, reconstructed as a muon with the same charge as the pion, a pion mass hypothesis. The invariant mass of the dimuon system can then be examined. The candidate is excluded if the dimuon mass (with the newly assigned mass hypotheses) is between 3036 MeV/ $c^2$  and 3156 MeV/ $c^2$ , *i.e.*

around the nominal  $J/\psi$  mass, and the pion fulfills  $\text{DLL}_{\mu\pi} > 5$  or satisfies the isMuon criterion. A similar procedure was implemented for the case where a kaon and a muon were swapped.

Further peaking backgrounds were rejected following the same procedure. These are  $B_s^0 \rightarrow \phi\mu^+\mu^-$ , where a kaon from the  $\phi$  is identified as a pion and  $B^+ \rightarrow K^+\mu^+\mu^-$ , where a soft pion is added from the rest of the event. Also considered were events of  $B^0 \rightarrow K^{*0}\mu^+\mu^-$ , where the daughter particles of the  $K^{*0}$  were swapped, leading to a wrong tag of the decay, and  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  where either the proton is misidentified as a pion or a kaon (and the kaon as a pion). The rejection proceeded similarly to the other peaking backgrounds.

Peaking backgrounds whose contribution were calculated to be negligible were  $B^+ \rightarrow K^{*+}\mu^+\mu^-$ , where a random kaon or pion is added (and one decay particle of the  $K^{*+}$  is not reconstructed);  $\Lambda_b^0 \rightarrow \Lambda^*(1520)\mu^+\mu^-$  with  $\Lambda^*(1520) \rightarrow pK^-$ , where the proton is misidentified as a pion;  $B^0 \rightarrow \rho\mu^+\mu^-$ , where a pion is misidentified as a kaon;  $B_s^0 \rightarrow f_0(980)\mu^+\mu^-$ , where a pion is misidentified as a kaon and  $B_s^0 \rightarrow K^{*0}\mu^+\mu^-$  without misidentification, but the wrong mother particle.

Another type of potential background is  $b \rightarrow c\mu$ ,  $c \rightarrow d\mu$  cascade decays, for example  $B^0 \rightarrow D^-\mu^+\nu_\mu$  with  $D^- \rightarrow K^{*0}\mu^-\bar{\nu}_\mu$ . These types of backgrounds can be largely suppressed by only accepting  $B^0$  candidates with a mass larger than  $5150 \text{ MeV}/c^2$ . Cascade decays tend to have a non-uniform angular distribution of the observables of interest for  $B^0 \rightarrow K^{*0}\mu^+\mu^-$ , which could bias the final result if they were present. It was therefore checked that the angular distributions of background events in the lower and the upper mass sidebands are compatible with each other, see Fig. 6.2.

The overall fraction of peaking backgrounds and the signal loss from the application of the vetoes are summarised in Table 6.3. All the numbers were estimated from simulation and using the branching fractions from the PDG [129].

Table 6.3.: Contamination of the final sample by specific peaking backgrounds and amount of signal loss due to the application of the vetoes for the peaking backgrounds. The numbers were calculated using simulation.

Decay	Background level [%]	Signal loss [%]
$B^0 \rightarrow K^{*0}\mu^+\mu^- (\pi \leftrightarrow K)$	$0.85 \pm 0.02$	0.11
$B^0 \rightarrow J/\psi K^{*0} (\pi \leftrightarrow \mu)$	$0.27 \pm 0.08$	0.05
$B^0 \rightarrow J/\psi K^{*0} (K \leftrightarrow \mu)$	$0.00 \pm 0.00$	0.03
$B_s^0 \rightarrow \phi\mu^+\mu^-$	$1.23 \pm 0.50$	0.32
$B^+ \rightarrow K^+\mu^+\mu^-$	$0.14 \pm 0.03$	0.00
$\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$	$0.75 \pm 0.15$	0.47
Total	$3.24 \pm 0.53$	0.98

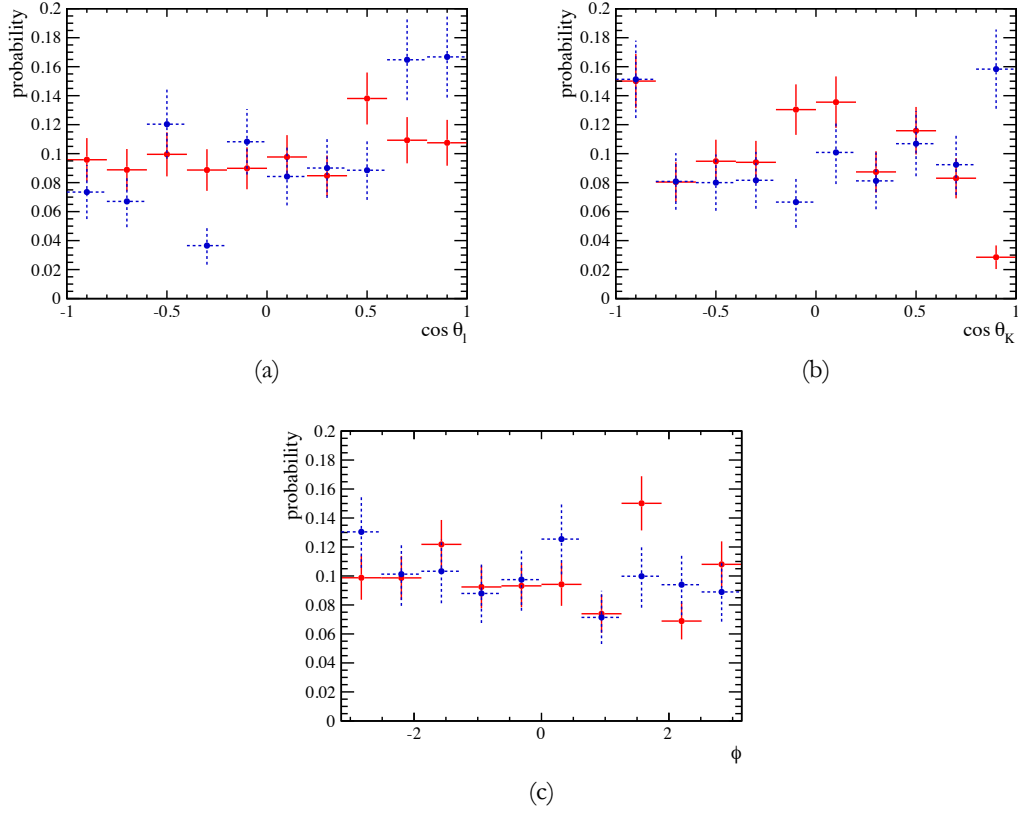


Figure 6.2.: Comparison of the angular distributions of the left (blue dashed) and the right (red) sideband for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  for  $\cos \theta_\ell$  (a),  $\cos \theta_K$  (b) and  $\phi$  (c).

## 6.2. Corrections to the simulation

The simulation is used to estimate the amount of peaking background (as seen in the last section) and to correct for detector acceptance effects. For this purpose, the distributions of the variables in the simulation must correspond to the ones in the collision data. Two different type of corrections to the simulated samples have been applied. If the variable is used in the BDT, it is adjusted or replaced before the application of the BDT. These variables include the impact parameter of the  $B^0$  and the four final state particles, the  $\text{DLL}_{K\pi}$  variable for particle identification of the  $K^+$  and  $\pi^-$  and the  $\text{DLL}_{\mu\pi}$  for particle identification of the muons. Furthermore the simulation is reweighted to account for different distributions of `isMuon`, which is used in the preselection, and the tracking efficiency. If a variable is not used in the BDT, the variable is reweighted so that the distributions in simulation and collision data agree. The different corrections are discussed below.

### 6.2.1. Impact parameter resolution

The impact parameter resolution in the  $x$ -coordinate in collision data and simulation is shown as a function of the inverse transverse momentum in Fig. 6.3. A clear difference between the two samples can be seen. The origin of this discrepancy is not fully understood. It is assumed to come from an incorrect material description in the RF foil in the VELO in the simulation. To adapt the impact parameter resolution of the particles, the track states are «smeared», *i.e.* a random number obtained from a Gaussian distribution is added to the  $x$  and  $y$ -coordinates of the state. This is done before the stripping selection is applied to the simulated sample. The effect of this smearing on the  $B^0$  impact parameter of  $B^0 \rightarrow J/\psi K^{*0}$  decays is shown in Fig. 6.4 and more information can be found in Ref. [132].

### 6.2.2. Muon identification

Muons are selected with the `isMuon` criterion and the  $\text{DLL}_{\mu\pi}$  classifier. DLL stands for «delta log likelihood» and is given by the following expression:

$$\text{DLL}_{\mu\pi} = \log(\mathcal{L}_\mu) - \log(\mathcal{L}_\pi), \quad (6.1)$$

where  $\mathcal{L}_\mu$  is the likelihood of the particle being a muon and  $\mathcal{L}_\pi$  the likelihood of the particle being a pion. The later is just used as a normalisation, as per definition every particle is considered a pion if no particle identification criteria are applied. The  $\text{DLL}_{\mu\pi}$  classifier therefore gives the logarithmic likelihood of a particle to be muon-like instead of pion-like: particles with negative values for  $\text{DLL}_{\mu\pi}$  are more pion-like than muon-like and vice versa for positive values. The  $\text{DLL}_{\mu\pi}$  classifier is built out of information of all subdetectors used for particle identification, but the main contribution for  $\text{DLL}_{\mu\pi}$  is provided by the muon system. While the basic principle is the same for all the other DLL classifiers, like

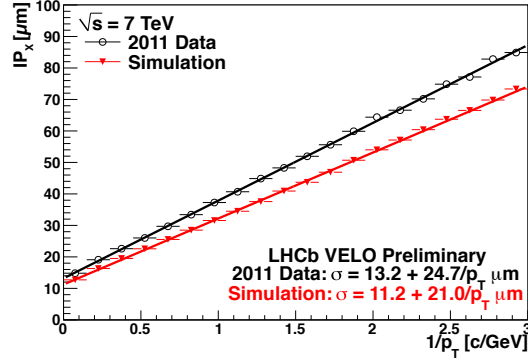


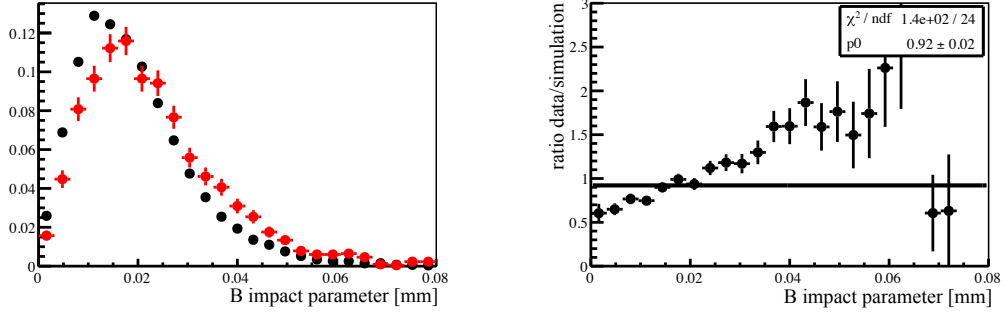
Figure 6.3.: Impact parameter resolution as a function of  $1/p_T$  in simulation and collision data taken in 2011. It can clearly be seen that the resolution is overestimated in the simulation compared to collision data. Figure taken from Ref. [52]

$DLL_{K\pi}$ , the contributing subdetectors have different weights depending on the specific classifier. For  $DLL_{K\pi}$ , the RICH detectors play the main role.

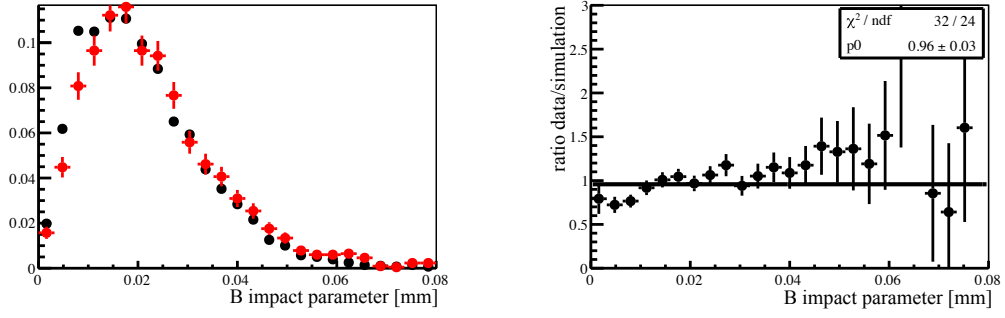
In addition to applying  $DLL_{\mu\pi}$ ,  $DLL_{K\pi}$  is applied to all final state particles. In the muon case it is used to reject hadrons misidentified as muons. The distributions of the classifiers  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  do not agree on collision data and on simulation. However, the disagreement for  $DLL_{\mu\pi}$  is smaller compared to the one for  $DLL_{K\pi}$ . The distribution of  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  for muons for collision data and simulation is shown in Fig. 6.5. The distribution is plotted using a tag-and-probe approach for  $J/\psi \rightarrow \mu^+\mu^-$  from the decay  $B^+ \rightarrow J/\psi K^+$ , where hard cuts are applied to the tag-muon so that the dimuon invariant mass distribution is background free. Furthermore, the probe-muon is required not to have fired the trigger. Both muons were required to have passed the isMuon condition.

The strategy to correct the  $DLL_{K\pi}$  and  $DLL_{\mu\pi}$  on simulation was the following: First, the  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  distributions for real muons as a function of the momentum and pseudorapidity were determined. For this purpose, the decay  $B^+ \rightarrow J/\psi K^+$  with  $J/\psi \rightarrow \mu^+\mu^-$  and a tag-and-probe approach is used. The binning in the momentum and the pseudorapidity was chosen as the largest deviation is expected between collision data and simulation in these two variables. Furthermore, these two variables have different distributions for  $B^+ \rightarrow J/\psi K^+$  and  $B^0 \rightarrow K^{*0}\mu^+\mu^-$ . The selection of the muons and the calculation of the classifier-values are explained below. For each bin in momentum and  $\eta$ , the  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  distributions are determined and a new  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  value for the muons in the simulation is then assigned according to the distributions.

The stripping selection for the decay  $B^+ \rightarrow J/\psi K^+$ , developed especially for particle identification studies, first selects oppositely charged dimuon pairs with a tag-and-probe approach. The dimuon system is then combined with a kaon to form  $B^+$  candidate. The



(a)



(b)

Figure 6.4.: Comparison between the impact parameter of the  $B^0$  in  $B^0 \rightarrow J/\psi K^{*0}$  decays in 2010 collision data (red) and simulation (black), before the smearing (a) and after the smearing (b). Figure taken from Ref. [132].

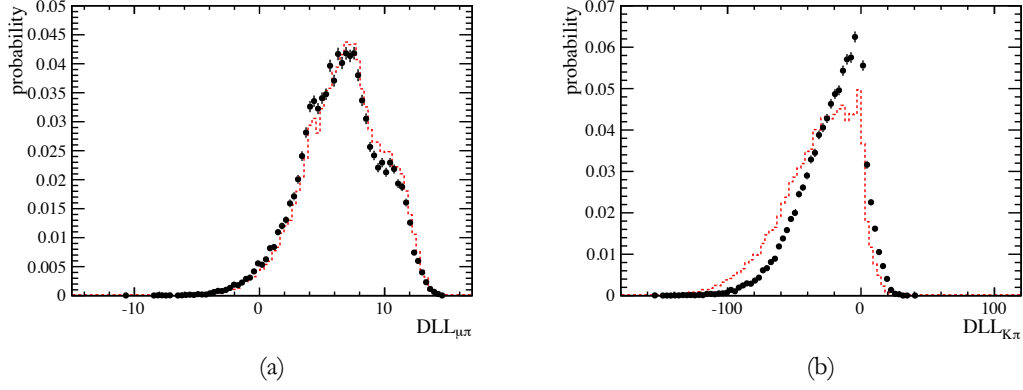


Figure 6.5.: Comparison between the  $DLL_{\mu\pi}$  (a) and  $DLL_{K\pi}$  (b) classifiers on real muons from  $J/\psi \rightarrow \mu^+\mu^-$ , coming from  $B^+ \rightarrow J/\psi K^+$ . In red is simulation and in black collision data. The isMuon condition is applied in these plots.

additional constraint on the  $B^+$  mass improves the background rejection. The stripping selection cuts are given in Table 6.4.

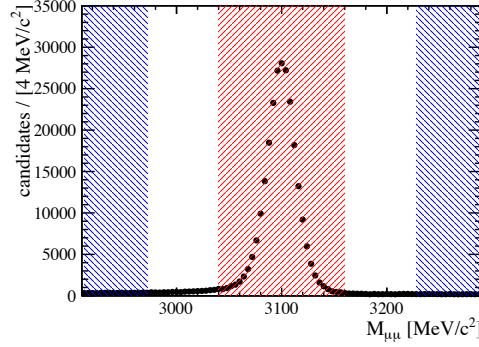


Figure 6.6.: Dimuon mass distribution when selecting the decay  $B^+ \rightarrow J/\psi K^+$ . In blue are the sidebands and in red is the signal region to determine the isMuon,  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$  efficiency.

More cuts are applied to further clean the dimuon invariant mass distribution; they are summarised in Table 6.5. In addition, the event was required not to be triggered exclusively on the probe-muon.

The dimuon invariant mass is divided into three regions to allow a statistical subtraction of potential background. The region between  $3040 \text{ MeV}/c^2$  and  $3160 \text{ MeV}/c^2$  is the signal region, the regions of  $2912 \text{ MeV}/c^2 - 2972 \text{ MeV}/c^2$  and  $3228 \text{ MeV}/c^2 - 3288 \text{ MeV}/c^2$  are the

sideband regions. The gap between the signal and the background region was to avoid contamination from the radiative tail and also take into account that for high momentum tracks the  $J/\psi$  peak is broader than for low momentum tracks. The dimuon invariant mass and the different regions are shown in Fig. 6.6.

The efficiency for a given cut on  $\text{DLL}_{\mu\pi}$  can then be calculated in the following way:

$$\varepsilon_{\text{DLL}_{\mu\pi}} = \frac{\# \text{ of events with } \text{DLL}_{\mu\pi} > \text{cutValue on the tag-muon}}{\# \text{ of events}}, \quad (6.2)$$

where the «# of events» is the number of events in the signal region minus the number of events in the two sidebands. This is done for every bin in momentum and pseudorapidity, which allows the construction of the  $\text{DLL}_{\mu\pi}$  distribution for each of these bins. A random value of  $\text{DLL}_{\mu\pi}$  is drawn according to this distribution in the corresponding momentum and pseudorapidity bin of the simulated particle. The  $\text{DLL}_{\mu\pi}$  value of the simulated particle is then replaced with this random value. The same procedure is done for  $\text{DLL}_{K\pi}$  for muons.

The uncertainty on the efficiency is given by:

$$\sigma_{\varepsilon_{\text{DLL}_{\mu\pi}}} = \sqrt{\frac{(1 - 2\varepsilon)(k_{\text{sig}} + k_{\text{SB}}) + \varepsilon^2(N_{\text{sig}} + N_{\text{SB}})}{(N_{\text{sig}} - N_{\text{SB}})^2}}, \quad (6.3)$$

where  $k_{\text{sig}}$  and  $k_{\text{SB}}$  correspond to the numerator in Eq. 6.2 for the signal region and the sidebands, respectively, and  $N_{\text{sig}}$  and  $N_{\text{SB}}$  correspond to the denominator for signal region and sidebands. The formula is derived in Appendix B. Note that this formula slightly overestimates the uncertainty as it also assigns an uncertainty on the total number of signal events. This in principle should not be done in a tag-and-probe approach: the numerator in Eq. 6.2 is a strict subset of the denominator. However, this small «conservatism» was not considered to be a problem.

A similar method was used to obtain the efficiency for isMuon on data. The same stripping and offline selections were used as for the determination of  $\text{DLL}_{\mu\pi}$  and  $\text{DLL}_{K\pi}$ , however without applying the isMuon condition on the tag-muon. The efficiency can then be determined the same way as described for  $\text{DLL}_{\mu\pi}$ , the only difference being the fact that only one cut value exists, as isMuon is Boolean. The efficiency for each momentum and pseudorapidity bin to pass isMuon was used to weight the simulated events with this efficiency. The efficiency map for isMuon and the ratio between collision data and simulation is shown in Fig. 6.7.

### 6.2.3. Systematic uncertainty for the muon identification

To check for systematic effects arising from the procedure to obtain the  $\text{DLL}_{\mu\pi}$  and  $\text{DLL}_{K\pi}$  values for muons, the following method was adopted. The  $B^+ \rightarrow J/\psi K^+$  sample from collision data was split in two halves. The first half was used to determine the  $\text{DLL}_{\mu\pi}$  distribution on collision data, the second half was used to compare the effect of applying a



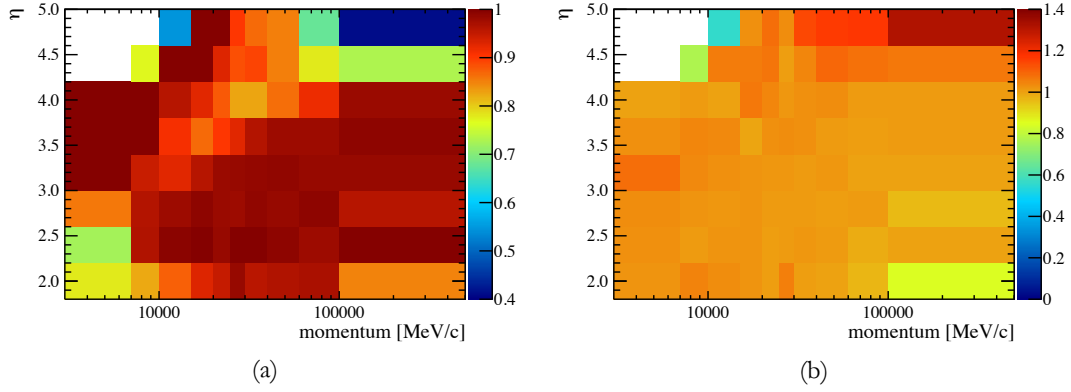


Figure 6.7.: Efficiency maps for the isMuon classifier. (a) shows the efficiency for passing the isMuon criterion as a function of the momentum and pseudorapidity in collision data, determined with the tag-and-probe method on  $B^+ \rightarrow J/\psi K^+$ . (b) shows the ratio of the efficiencies in collision data and simulation, where both efficiencies were determined on  $B^+ \rightarrow J/\psi K^+$ .

Table 6.4.: Selection cuts for  $B^+ \rightarrow J/\psi K^+$  in the stripping selection.

Particle	Variable	Value
	isMuon	true
tag-muon	$p$	$> 6 \text{ GeV}/c$
	$p_T$	$> 1.5 \text{ GeV}/c$
	$\chi_{\text{IP}}^2$	$> 25$
$K^+$	$\chi_{\text{IP}}^2$	$> 25$
$\mu^+ \mu^-$	$\chi_{\text{vtx}}^2/\text{ndof}$	$< 8$
	$\chi_{\text{FD}}^2$	$> 225$
$\mu^+ \mu^- K^+$	$\chi_{\text{IP}}^2$	$< 25$

Table 6.5.: Selection cuts for the  $B^+ \rightarrow J/\psi K^+$  offline selection for particle identification studies.

Particle	Variable	Value
tag-muon	isMuon	true
$\bar{K}^+$	$DLL_{K\pi}$	$> 0$
	DIRA	$> 0.999$
$\mu^+ \mu^- K^+$	mass	$< 3160 \text{ MeV}/c^2$ and $> 2940 \text{ MeV}/c^2$
	# of candidates/event	1

cut on the negative muon from  $B^+ \rightarrow J/\psi K^+$  to weighting the event with the efficiency corresponding to the  $DLL_{\mu\pi}$  cut for the negative muon. The number of events that pass the cut, or are weighted, was then determined with a fit to the dimuon distribution where a double Gaussian<sup>7</sup> was used to describe the signal distribution and an exponential function to describe the background distribution. The same procedure was followed for  $DLL_{K\pi}$ . The result of this comparison can be seen in Fig. 6.8 and Fig. 6.9. Both distributions show good agreement for regions of high efficiency, but worse agreement for low efficiency of the  $DLL_{\mu\pi}$  or  $DLL_{K\pi}$  cut. Although there is no fixed cut value for  $DLL_{\mu\pi}$  and  $DLL_{K\pi}$ , as both variables are included in a BDT, preferentially values corresponding to high efficiencies will be used. The growing discrepancies when moving to higher values of the classifiers are, therefore, not significant. Furthermore it should be noted that the statistics for high values was too low to determine the systematic effect accurately.

It is not entirely clear where the (small) initial discrepancy comes from. It could be due to background which is not totally flat. Therefore the sideband subtraction would not exactly reproduce the number of signal events. Moreover it is possible that there are tracks outside the region which is covered in the momentum and pseudorapidity bins. However, for the analysis of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  these effects and their resulting discrepancy are all without relevance.

#### 6.2.4. Kaon identification

Kaons and pions are identified using the  $DLL_{K\pi}$  variable whose distribution is very different in simulation and collision data. The following procedure, very similar to the one for muons, was therefore adopted: a sample of genuine pions and kaons was selected in colli-

<sup>7</sup>A double Gaussian is the sum of two Gaussian PDFs. For the fit described in this section, the fraction between the two Gaussian functions was determined when fitting the dimuon mass distribution without cuts and then fixed to this value when determining the amount of signal for the case with the cuts and the weighting.

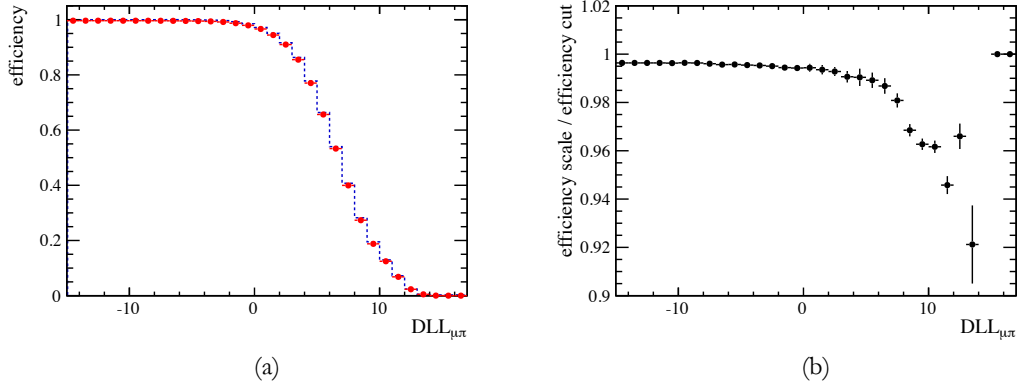


Figure 6.8.: In (a) is shown the comparison between cutting on the  $DLL_{\mu\pi}$  of the negative muon in  $B^+ \rightarrow J/\psi K^+$  (blue dashed line) and weighting the event according to the determined efficiency (red data points). In (b) is shown the ratio of both efficiencies. Note that the errors are strongly correlated between the individual bins and that the values and errors are not accurate anymore above a value of about 5.

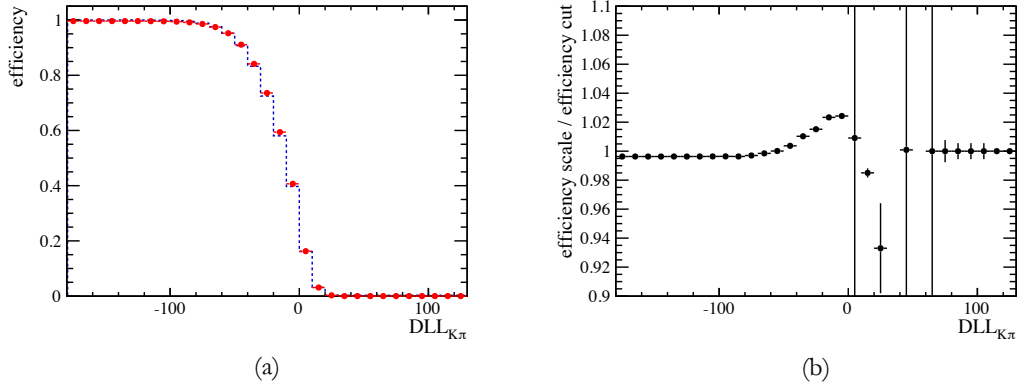


Figure 6.9.: In (a) is shown the comparison between cutting on the  $DLL_{K\pi}$  of the negative muon in  $B^+ \rightarrow J/\psi K^+$  (blue dashed line) and weighting the event according to the determined efficiency (red data points). In (b) is shown the ratio of both efficiencies. Note that the errors are strongly correlated between the individual bins and that the values and errors are not accurate anymore above a value of about 0.

sion data, using the decays  $D^{*+} \rightarrow D^0 \pi^+$  with  $D^0 \rightarrow K^- \pi^+$ ,  $\Lambda \rightarrow p \pi^-$  and  $K_s^0 \rightarrow \pi^+ \pi^-$ . The final state particles can be selected without using particle identification. The  $DLL_{K\pi}$  variable was sampled for pions and kaons from these decays in momentum, pseudorapidity and number of tracks. Each pion and kaon in the simulation then gets a new  $DLL_{K\pi}$  value assigned, which is subsequently used in the selection.

Furthermore, the  $DLL_{\mu\pi}$  classifier was also applied to veto muons misidentified as kaons or pions. The procedure to extract this information is the same as explained before.

### 6.2.5. Further corrections and checks to the simulation

**Tracking efficiency:** To account for differences in the track reconstruction efficiency between simulation and collision data, the simulated events are weighted according to the two-dimensional efficiency map as presented in Sect. 4.8.

**Kinematic differences:** The momentum and transverse momentum spectra of  $B^0 \rightarrow J/\psi K^{*0}$  were compared in collision data and simulation and found to be in good agreement. No perfect agreement is expected, as in the collision data, a pollution by an  $S$ -wave<sup>8</sup> the component is present, which is not the case in the simulation.

The output of the BDT after the application of the trigger, stripping and the preselection for  $B^0 \rightarrow J/\psi K^{*0}$  is shown in Fig. 6.10. It shows good agreement between simulation and collision data for the control channel after the application of the previously discussed corrections.

## 6.3. Correction for detector acceptance

One of the main goals in the analysis of the decay of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is the extraction of physics parameters by an angular analysis. However, there is a complication: not only the underlying physics can influence the angular distributions of the final state particles, but also the detector acceptance. Detector acceptance in this context means the explicit detector acceptance given by the LHCb geometry, but also the implicit acceptance given by, for example, cuts that tend to favour tracks in one direction over another direction. Examples are the cuts on the impact parameter, which disfavours events which are very forward, as they tend to have a smaller impact parameter, or the requirement that a muon must have a momentum greater than 3 GeV/c in order to leave enough information in the muon stations to be identified.

An acceptance correction was developed to correct for these distortions of the angular distributions. A large sample of simulated  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  events was generated where

---

<sup>8</sup>The decay  $B^0 \rightarrow J/\psi K^{*0}$  has the same final state as the decay  $B^0 \rightarrow J/\psi K^+ \pi^-$ , *i.e.* without the intermediate  $K^{*0}$  resonance. These two decays cannot be disentangled in the invariant mass, as the non-resonant contribution is always present in the  $K^{*0}$  mass region. However, the non-resonant decay proceeds via an  $S$ -wave while the resonant decay proceeds via a  $P$ -wave.

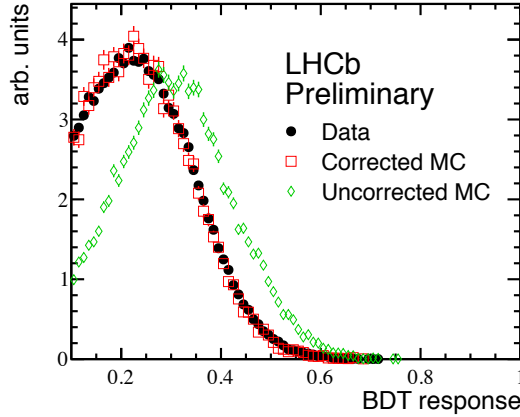


Figure 6.10.: Comparison between the BDT outputs for collision data (black), uncorrected simulation (green) and corrected simulation (red) on the control channel  $B^0 \rightarrow J/\psi K^{*0}$  after the application of the trigger, stripping and offline pre-selection. The effect of correcting the simulated distributions to the ones from collision data can clearly be seen.

the decay of the particles was only steered by the available phase space. This results in flat distributions in the angles and a falling distribution in  $q^2$ . The simulated events were then corrected for the differences between simulation and collision data as described in the previous section. Then they were required to pass the trigger, the stripping and all the offline selections. A transformation is applied to maximise the yield from the available sample: the acceptance is assumed to be symmetric in  $\cos \theta_l$  about  $0^9$  and in  $\phi$  if the transformation  $\phi \rightarrow \phi + \pi$  is applied. A weight can then be used for every event in a maximum likelihood fit. In the subsequent analysis. This weight corresponds to the inverse efficiency of an event to pass all the described stages.<sup>10</sup>

Another possibility would be to incorporate the acceptance function in the PDF which is then fitted to data. However, the acceptance correction has to be done in bins which are more narrow in  $q^2$  than those which are used for the analysis (see later). Therefore, this possibility is not used.

The most general approach to the acceptance correction would be to bin in all three angular variables,  $\cos \theta_l$ ,  $\cos \theta_K$ ,  $\phi$ , and  $q^2$  and then derive the correction for every four dimensional bin individually. However, this would require large amounts of simulated data. A different approach was therefore chosen. For a given bin in  $q^2$ , the efficiency can be factorised for the three angular variables. The distribution of these variables is then fitted with a

<sup>9</sup>This assumption would only fail if there was a large difference in efficiency for  $\mu^+$  and  $\mu^-$  in the full sample and a large  $\mathcal{CP}$  asymmetry between  $B^0$  and  $\bar{B}^0$ .

<sup>10</sup>The weight is normalised such that its mean is one.

6<sup>th</sup> order polynomial. It was checked that this factorised approach reproduces the distribution of phase space events after the trigger, the reconstruction and the selection were applied. Furthermore the events were weighted with their inverse efficiency and compared to the distribution of the phase space generated events without applying the trigger, the reconstruction or the selection. Both distributions agree with each other.

## 7. Results of the «first» angular analysis of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$

What is referred to as the «first» angular analysis comprises the measurement of the variables  $A_{FB}$ ,  $F_L$ ,  $A_T^{(2)}$ ,  $A_T^{Re}$ ,  $S_3$ ,  $S_9$  and  $A_9$  with  $\approx 1 \text{ fb}^{-1}$  of collision data collected in 2011. The analysis is described in detail in Ref. [126]. The main results are shown in Fig. 7.1 and 7.2. No result is shown for the variable  $S_9$ ; its experimental results are compatible with zero for the full  $q^2$  range which corresponds to the SM prediction.

All observables are compatible with their Standard Model predictions in the full  $q^2$  range.

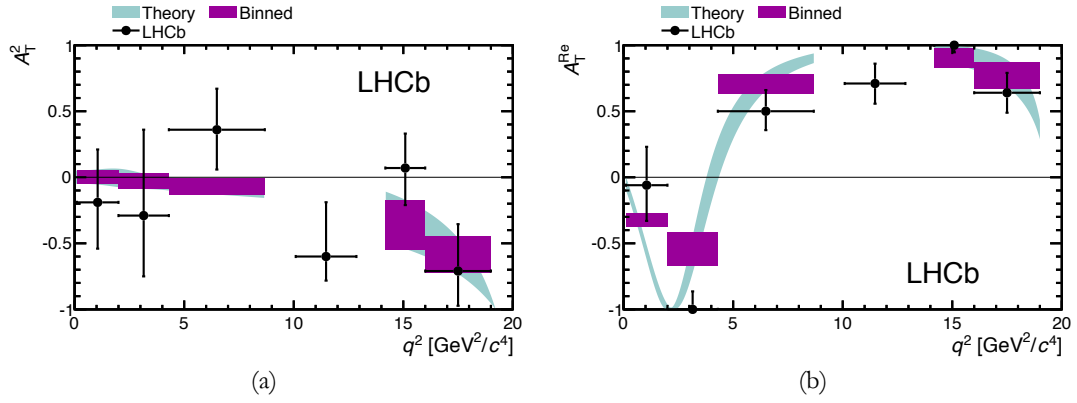


Figure 7.1.: Measurement of the observables  $A_T^{(2)}$  (a) and  $A_T^{Re}$  (b) which have reduced form factor uncertainties. Both observables are compatible with the Standard Model in the full  $q^2$  range. The procedure to extract these observables and the origin of the SM prediction is given in Ref. [126].

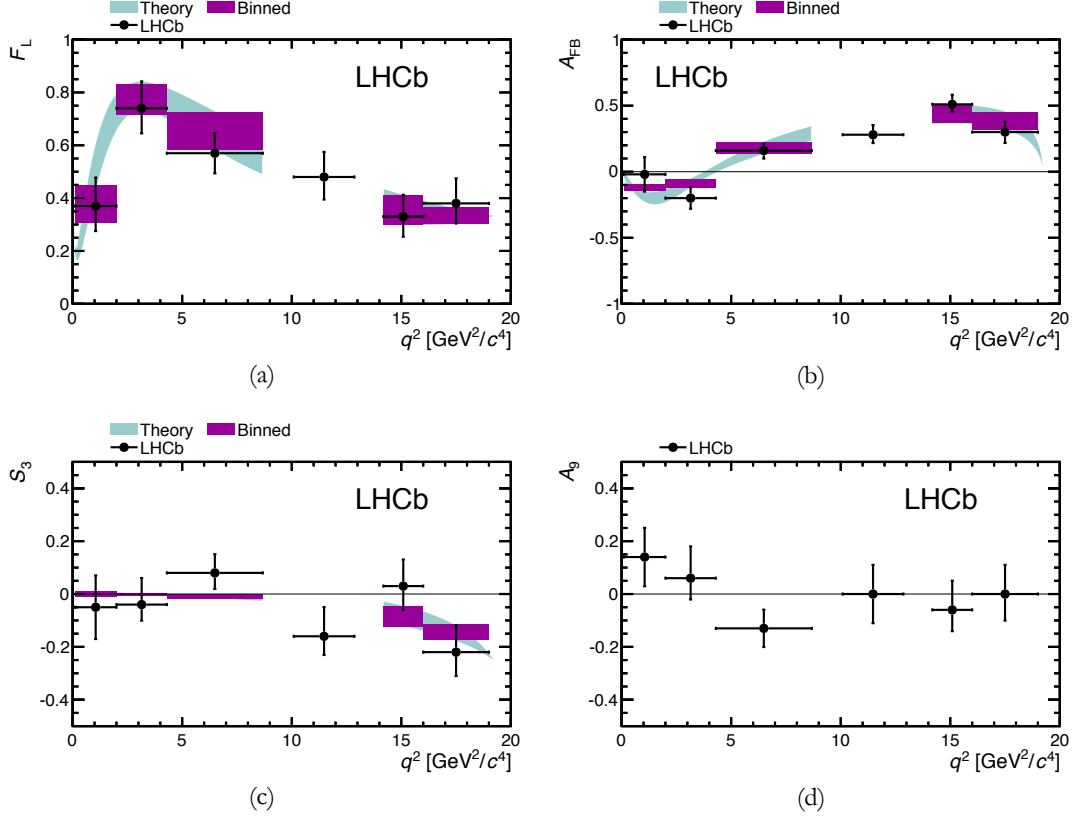


Figure 7.2.: Measurement of the observables  $F_L$  (a),  $A_{FB}$  (b),  $S_3$  (c) and  $A_9$  (d). For  $A_9$  the SM value is vanishingly small. All observables are compatible with the Standard Model in the full  $q^2$  range. The procedure to extract these observables and the origin of the SM prediction is given in Ref. [126].



## 8. Determination of the zero-crossing point of $A_{\text{FB}}$

The prediction for the zero-crossing point of  $A_{\text{FB}}(q^2)$  in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is particularly clean as the form-factors and, therefore, their uncertainties cancel to first order. The measurement presented here is the first of its kind: the zero-crossing point was not measured previously in LHCb or in any other experiment. It requires a certain amount of accumulated data to make a measurement with a reasonable statistical uncertainty. Furthermore some precautions have to be taken in order not to bias the measured value of the zero-crossing point with the experimental method.

This chapter gives an overview of the experimental procedure to extract the zero-crossing point, including crosschecks for the procedure itself and an estimate of the statistical and systematical uncertainties.

### 8.1. Extraction of the zero-crossing point with a linear fit

The simplest method to determine the zero-crossing point of  $A_{\text{FB}}$  is to do a binned, linear fit of  $A_{\text{FB}}(q^2)$  in a region around the zero-crossing [133]. The idea is to split the number of events in two samples, forward events with  $\cos \theta_\ell > 0$  and backward events with  $\cos \theta_\ell < 0$ , and count the number of signal events. The forward-backward asymmetry  $A_{\text{FB}}$  can then be formed as:

$$A_{\text{FB}} = \frac{\# \text{ forward signal events} - \# \text{ backward signal events}}{\# \text{ forward signal events} + \# \text{ backward signal events}} \quad (8.1)$$

for each bin in  $q^2$ .

However, this method has several drawbacks. First of all, it cannot be applied «blindly»<sup>1</sup>, as the approximate shape of  $A_{\text{FB}}$  has to be known to define a fit region. Second, the choice of the fit region can severely bias the result for the zero-crossing point. And lastly, it implicitly assumes that the shape of  $A_{\text{FB}}$  is straight enough to be fitted with a straight line. This is the case for  $A_{\text{FB}}$  in the Standard Model but need not necessarily be the case for all physics models.

---

<sup>1</sup>i.e. without looking at the  $A_{\text{FB}}$  distribution.

To check biases induced by these effects, a study was done with simulated  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  events without background. Two binning schemes were chosen: One from 0 to 8  $\text{GeV}^2/c^4$  with a bin width of 1  $\text{GeV}^2/c^4$ , the other from 0 to 8  $\text{GeV}^2/c^4$  with a bin width of 0.5  $\text{GeV}^2/c^4$ . The number of events in each individual bin was calculated by assuming that in total 1000  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  signal events are reconstructed in  $1 \text{ fb}^{-1}$  for the full  $q^2$  range. This corresponds (approximately) to the size of the 2011 data sample. «Toy samples»<sup>2</sup> were then generated for forward and backward events in each bin according to a Poisson distribution with the mean value taken from the simulated events.

The forward-backward asymmetry,  $A_{\text{FB}}$ , calculated following Eq. 8.1, was then fitted with a linear function:

$$A_{\text{FB}}(q^2) = a_0 \cdot (q^2 - q_0^2), \quad (8.2)$$

in the range  $[q_{\text{low}}^2, q_{\text{high}}^2]$  where  $q_0^2$  is the zero crossing point and  $a_0$  a scaling parameter. To check for possible biases, the residual of the zero-crossing points in the individual toy samples with respect to the zero-crossing point in the simulated data was plotted. The results for different  $q_{\text{low}}^2$  and  $q_{\text{high}}^2$  are shown in Fig. 8.1. It is clearly visible that shifting the start and end points by 1  $\text{GeV}^2/c^4$  introduces a large bias in the residual distributions, which is non-negligible with respect to the expected statistical uncertainty. Figure 8.2 shows two examples of toy samples with a bin width of 1  $\text{GeV}^2/c^4$  and 0.5  $\text{GeV}^2/c^4$  which illustrate the difficulty to choose a start and end point for the linear fit with  $1 \text{ fb}^{-1}$  of data. Therefore an uncertainty of 1  $\text{GeV}^2/c^4$  in the choice of the start or end point, as illustrated in Fig. 8.1, is a realistic case when considering the collision data.

As the zero-crossing point is not a parameter of the  $A_{\text{FB}}$  distribution of the Standard Model, it had to be determined from the simulated sample itself to plot the residuals. This was accomplished by using the full simulated sample and splitting the range from 0  $\text{GeV}^2/c^4$  and 6  $\text{GeV}^2/c^4$  into 50 bins. For each bin  $A_{\text{FB}}$  was calculated and the distribution fitted with a linear fit from 3  $\text{GeV}^2/c^4$  to 5  $\text{GeV}^2/c^4$ . This is shown in Appendix C. The extracted zero-crossing point is 4.07  $\text{GeV}^2/c^4$  with an uncertainty of less than 1%.

*In somma*, the linear fit method cannot be used to determine the zero-crossing point of  $A_{\text{FB}}$  in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  with  $1 \text{ fb}^{-1}$  of collision data as it introduces significant biases.

## 8.2. The «unbinned counting» technique

To overcome the limits of the linear fit, a different approach, called «unbinned counting» was chosen [134]. The idea is the following: The events in a given range of  $q^2$  are split up in forward ( $\cos \theta_\ell > 0$ ) and backward ( $\cos \theta_\ell < 0$ ). The forward and backward distribution are then fitted with a high order polynomial in an extended unbinned maximum likelihood fit. The forward-backward asymmetry can then be calculated as:

<sup>2</sup>The prefix «toy» refers to simulations or simulated data, which was not created using the full simulation of LHCb but rather by generating events using PDFs as input and sampling the distributions using random number generators.

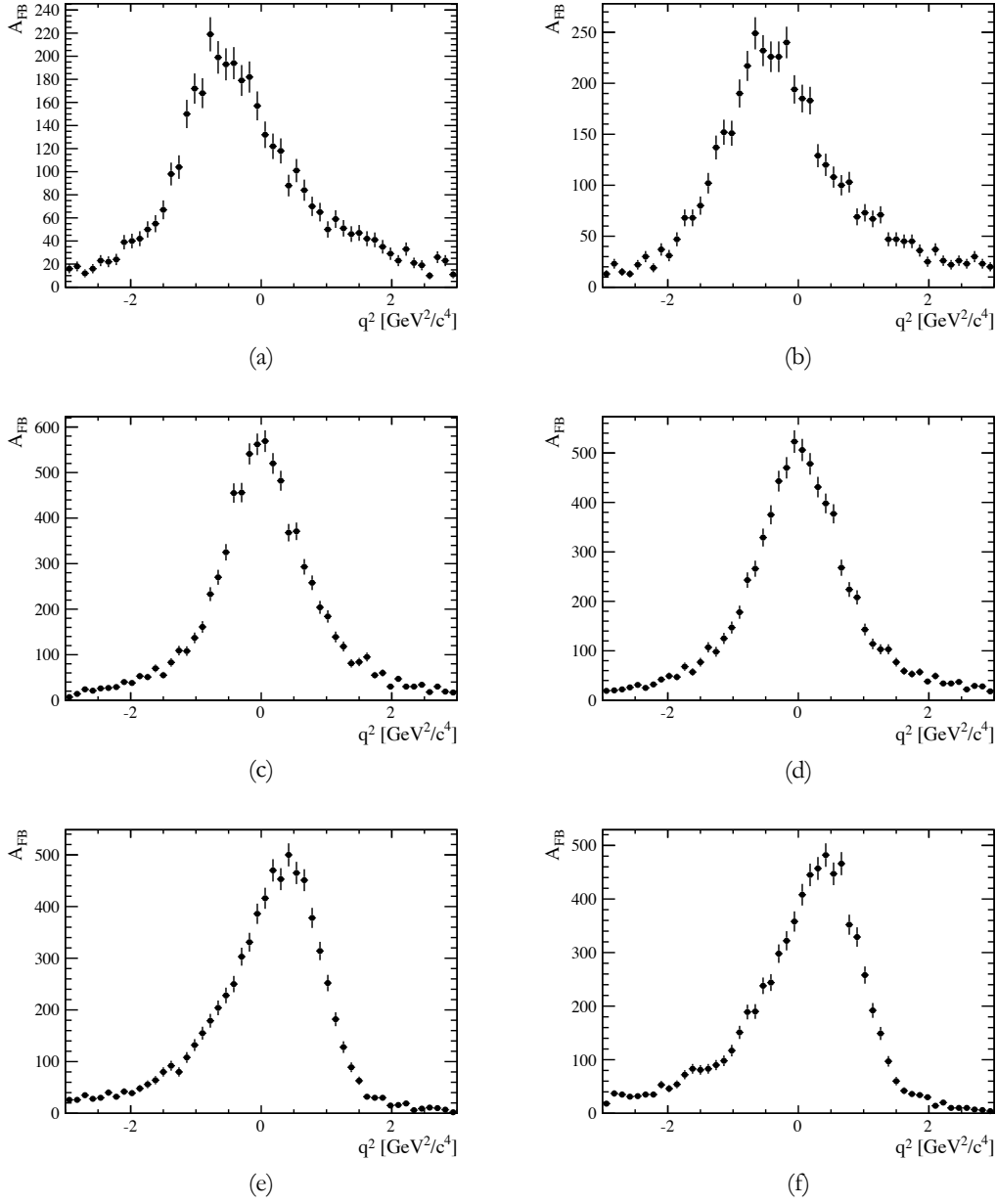


Figure 8.1.: Residuals of zero-crossing points in simulation of many toy samples with respect to the zero-crossing point of the full simulation sample when changing the range of the linear fit. (a), (c) and (e) show fits in the intervals  $1 - 5 \text{ GeV}^2/c^4$ ,  $2 - 6 \text{ GeV}^2/c^4$  and  $3 - 7 \text{ GeV}^2/c^4$  respectively, with a bin width of  $1 \text{ GeV}^2/c^4$ . (b), (d) and (f) show fits in the intervals  $1 - 5 \text{ GeV}^2/c^4$ ,  $2 - 6 \text{ GeV}^2/c^4$  and  $3 - 7 \text{ GeV}^2/c^4$  respectively, with a bin width of  $0.5 \text{ GeV}^2/c^4$ . A bias of the mean of the distribution of the first and the last row with respect to the second row is clearly visible.

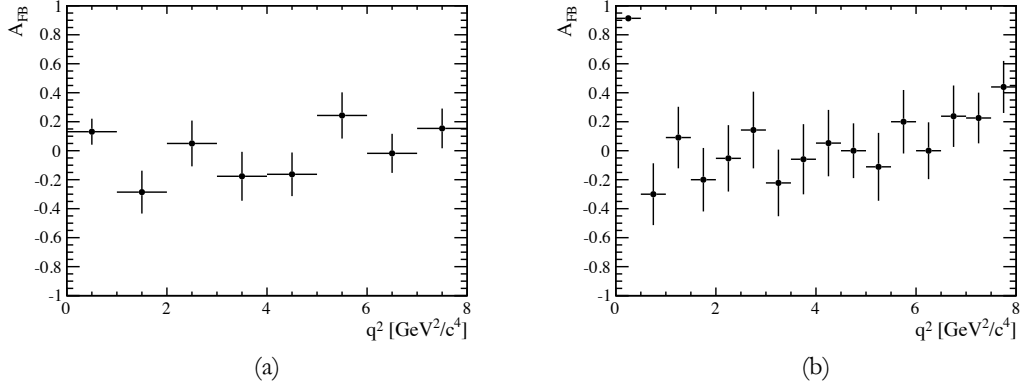


Figure 8.2.: Examples for  $A_{FB}$  distributions with a  $1 \text{ GeV}^2/c^4$  (a) and  $0.5 \text{ GeV}^2/c^4$  binning (b) scheme. They illustrate the difficulty to choose a reasonable start and end point of the linear fit to determine the zero-crossing point of  $A_{FB}$ .

$$A_{FB}(q^2) = \frac{n_F \cdot PDF_F(q^2) - n_B \cdot PDF_B(q^2)}{n_F \cdot PDF_F(q^2) + n_B \cdot PDF_B(q^2)}, \quad (8.3)$$

where  $n_F, n_B$  are the number of forward and backward events, respectively, and  $PDF_F, PDF_B$  the probability density functions that describe the forward and backward distribution, respectively. The zero-crossing point can easily be determined by finding the root of  $A_{FB}(q^2)$  with numerical recipes.

The name «unbinned counting» is strictly speaking a misnomer, as it is not different from an extended unbinned maximum likelihood fit. However, it was chosen to clearly state the difference between the angular measurements and the determination of the zero-crossing point of  $A_{FB}$ .

### 8.2.1. Choice of the order of the polynomial

As the analytical description of the forward and backward distributions in  $q^2$  strongly depends on form factors, no generic choice for the order of the polynomial exists. However, for a given sample of collision data, the goodness-of-fit can be tested for different orders. This is described in Ref. [127]. For all subsequent studies, a third order polynomial<sup>3</sup> was used to describe the signal distribution for forward and backward events.

<sup>3</sup>In practice, only Tschebyscheff polynomials were used instead of linear ones where large correlations between the coefficients can affect the fit quality.

### 8.2.2. Independence of fit range

It had to be shown that the measured zero-crossing point does not depend on the range between which the forward and backward distributions are fitted. At maximum, the fit range extends from  $1 \text{ GeV}^2/c^4$  to  $7.8 \text{ GeV}^2/c^4$  to not be affected by the photon-pole<sup>4</sup> at  $0 \text{ GeV}^2/c^4$  and the radiative tail of the  $J/\psi$  when approaching the region of the charmonium resonances. Figure 8.3 shows four different ranges and the residual distributions of the zero-crossing points of many toy samples with respect to the zero-crossing point determined from the full simulated sample. No or only a small bias can be seen in these distributions due to the different fit ranges.

## 8.3. Choice and determination of probability distribution functions for the fit

To determine the zero-crossing point, a fit to the forward- and backward distributions in  $q^2$  is performed. However, in presence of background, the  $q^2$  distribution will contain a mixture of signal and background events. To disentangle these two categories, a two-dimensional fit is implemented: one dimension being  $q^2$  and the other dimension being the invariant mass. Using a combined PDF with signal and background components, the two contributions can be (statistically) separated.

The mass component of the signal is extracted by fitting the invariant mass distribution of  $B^0 \rightarrow J/\psi K^{*0}$  between  $5150 \text{ MeV}/c^2$  and  $5750 \text{ MeV}/c^2$  with two Crystal Ball functions, both having the radiative tail on the left side. For this fit, the nominal  $B^0$  mass and the shape parameters  $\alpha$  and  $n$  are shared between the two Crystal Ball functions.  $\alpha$  denotes the decay constant of the exponential fall off and  $n$  the turning point between the Gaussian component and the exponential component. The widths of both Crystal Ball distributions are allowed to float independently in this fit. Additionally, a double Crystal Ball function is used to parametrise the contribution from the decay  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  with the same constraints as the first double Crystal Ball function. Furthermore the yield of the second double Crystal Ball function was fixed with respect to the first one and the shift of the nominal  $B_s^0$  mass value with respect to the  $B^0$  mass value was set constant as well.

For the fit to the  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  invariant mass distribution, only the PDF for  $B^0 \rightarrow J/\psi K^{*0}$  was used and all parameters were fixed to the values obtained in the fit to  $B^0 \rightarrow J/\psi K^{*0}$ . The only floating parameter was the number of events. The mass component of the background for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  was fitted with an exponential function, where the decay constant was allowed to float. For the  $q^2$  component of the signal, a third order polynomial was used (*cf.* Sect. 8.2.1) while for the background a second-order polynomial was used.

---

<sup>4</sup>For  $q^2 \rightarrow 0$  the branching fraction strongly increases which prevents fitting a third order polynomial.

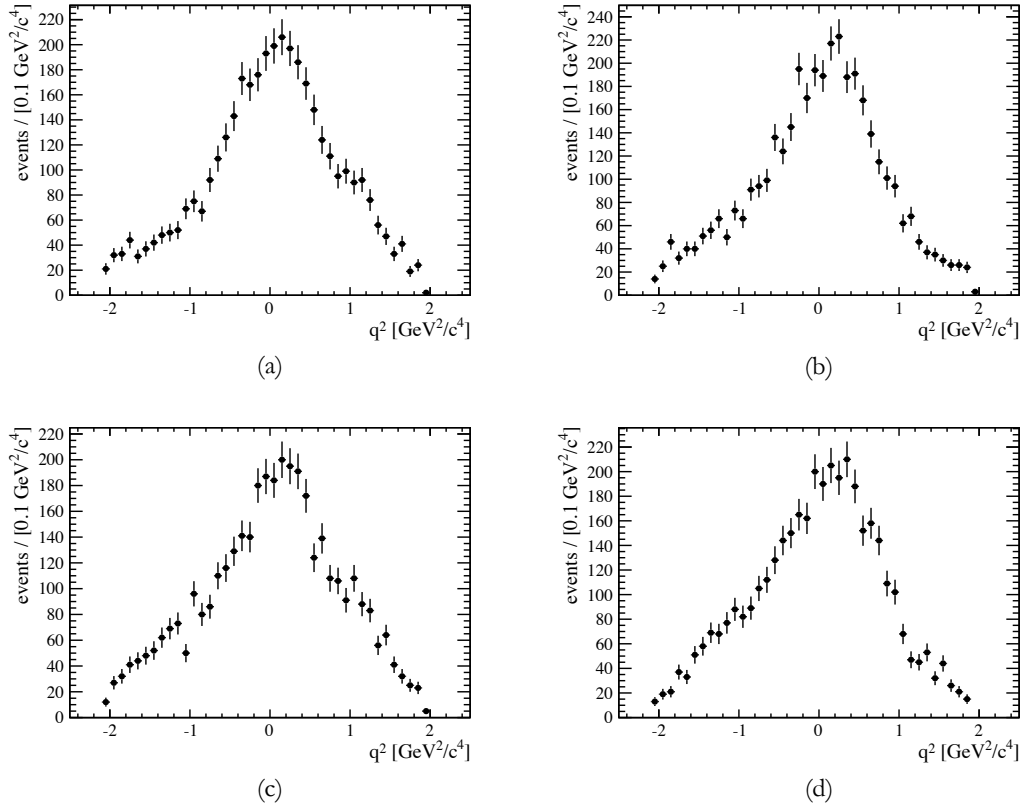


Figure 8.3.: Residuals of zero-crossing points in simulation of many toy samples with respect to the zero-crossing point of the full simulation sample when changing the range of the unbinned counting. Figures (a), (b), (c) and (d) show ranges between  $1 - 7 \text{ GeV}^2/c^4$ ,  $1 - 7.8 \text{ GeV}^2/c^4$ ,  $1.5 - 7.8 \text{ GeV}^2/c^4$  and  $1.5 - 7.8 \text{ GeV}^2/c^4$  respectively. No or only a small bias can be seen between the figures, which is negligible with respect to the statistical uncertainty.

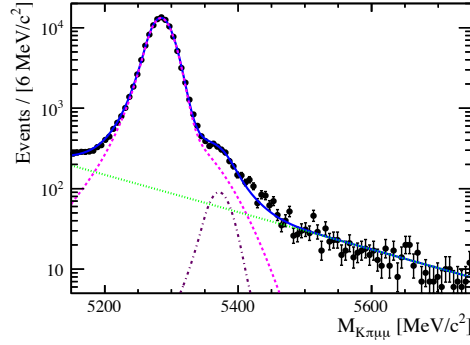


Figure 8.4.: Fit to the  $K\pi\mu\mu$  invariant mass distribution (solid blue) with contributions from  $B^0 \rightarrow J/\psi K^{*0}$  (dashed magenta),  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  (dashed dotted red) and combinatorial background (dotted green).

## 8.4. Checks on simulated data

### 8.4.1. Simulated data after reconstruction, signal only

The behaviour of the unbinned counting method was checked on simulated data corresponding to the Standard Model distribution implemented in the LHCb simulation program. In addition, a different set of simulated data, not corresponding to the Standard Model was used (called BSM model). This model has a zero-crossing point at  $3.43 \text{ GeV}^2/c^4$ . To check the behaviour, the mass distribution of the full simulated sample was fitted between  $5150 \text{ MeV}/c^2$  and  $5800 \text{ MeV}/c^2$  with the mass model described in Sect. 8.3<sup>5</sup>. The  $q^2$  distribution was fitted with a fourth order polynomial<sup>6</sup>. This two-dimensional PDF was then used to generate toy-samples where the total amount of forward- and backward events was fluctuated by Poisson distributions. The mean for the backward distribution was set to the number of backward events found in the collision data sample, the mean for the forward distribution was set to the number of backward events found in the collision data sample times the fraction of forward to backward events in the simulated sample. The mean was not set to the number of events in the forward distribution in order not to accidentally «unblind» the result of the zero-crossing point in these cross-checks. Each of the toy-samples was fitted with the same PDF as was used for their generation except that only a third order polynomial was used to fit the  $q^2$  distribution.  $A_{\text{FB}}$  was then calculated using

<sup>5</sup>The range in these studies is  $50 \text{ MeV}/c^2$  higher than in the fit to  $B^0 \rightarrow J/\psi K^{*0}$  to extract the signal shape and the final fit for the zero-crossing point. The reason is that the mass window changed over the course of the study and the simulation studies were already finished when the change occurred.

<sup>6</sup>While for all smaller samples, corresponding to the  $1 \text{ fb}^{-1}$  of collision data in 2011, a third order polynomial was sufficient, a fourth order polynomial was chosen for the large amount of simulated data to take small variations in  $q^2$  into account as well.

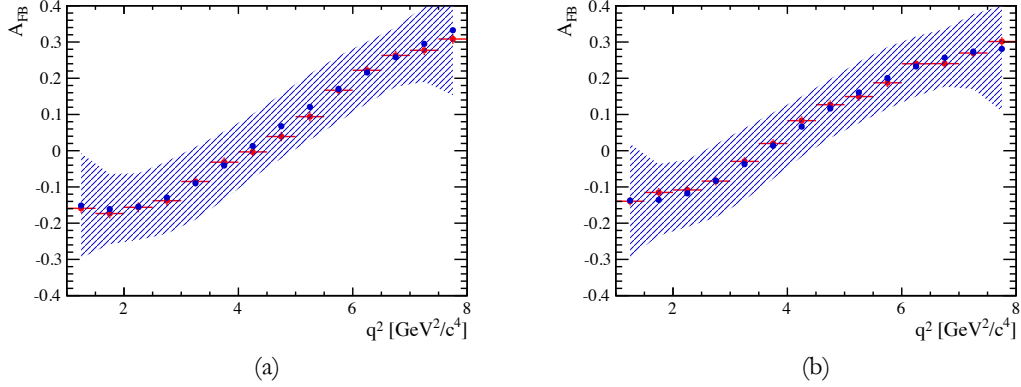


Figure 8.5.: Distribution of  $A_{\text{FB}}$  values obtained after the reconstruction using signal only for the Standard Model (a) and the BSM model (b). In blue are the mean values of  $A_{\text{FB}}$  of all toy experiments with the  $1\sigma$  error band. In red are the values as extracted from the simulated sample without reconstruction. The zero-crossing point in the Standard Model sample is  $4.07 \text{ GeV}^2/c^4$  while it is  $3.43 \text{ GeV}^2/c^4$  in the BSM model sample.

Eq. 8.3 for each sample. No acceptance correction was applied to any of these samples (neither for generating the toy-samples nor for fitting them).

In Fig. 8.5 the distribution of  $A_{\text{FB}}$  values for the Standard Model and the BSM model with their corresponding  $1\sigma$  error band can be seen. The mean of all  $A_{\text{FB}}$  values and the binned values, extracted from the simulated samples without applying the reconstruction, show a very good agreement. The differences are much smaller than the statistical uncertainty.

#### 8.4.2. Simulated data after reconstruction, signal and background

The same procedure as explained in the previous section was also performed after the addition of background events. The shape of the background in the mass and the  $q^2$  distributions was derived from a fit to both the lower and the upper sideband in collision data. The PDFs explained in Sect. 8.3 were used. Then, as for the signal only case, toy-samples were generated with Poisson distributions for the number of forward and backward events, where the mean of these Poisson distributions was fixed to the number of events in the sidebands. These background toy samples were then mixed with the ones for signal. For the fit to the full distributions, containing signal and background, the distributions explained in Sect. 8.3 were again used. The result of the fit is shown in Fig. 8.6. There is a good agreement between the mean of  $A_{\text{FB}}$  as derived from the individual toy samples and the value of the full simulated sample without reconstruction.



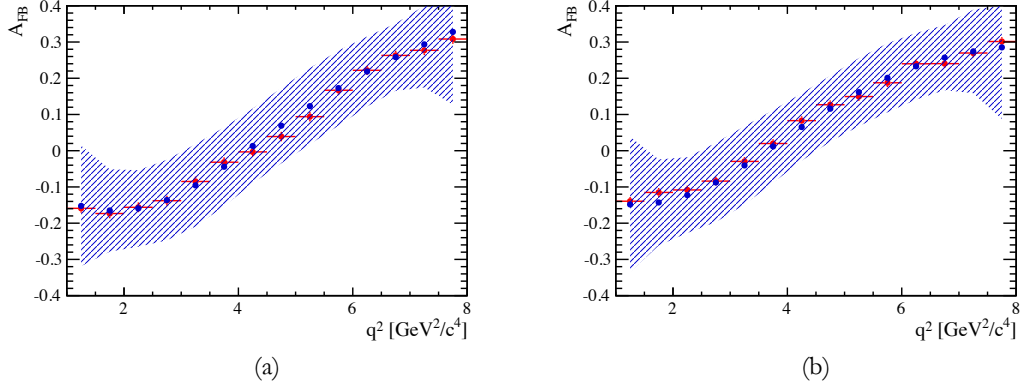


Figure 8.6.: Distribution of  $A_{\text{FB}}$  values obtained after the reconstruction using signal and background for the Standard Model (a) and the BSM model (b). In blue are the mean values of  $A_{\text{FB}}$  of all toy experiments with the  $1\sigma$  error band. In red are the values as extracted from the simulated sample without reconstruction. The zero-crossing point in the Standard Model sample is  $4.07 \text{ GeV}^2/c^4$  while it is  $3.43 \text{ GeV}^2/c^4$  in the BSM model sample.

### 8.4.3. Determination of the zero crossing point

Finding the zero-crossing point in a given sample is equivalent to finding the root of a curve. It is performed in the following way<sup>7</sup>: the range between a start and an end point in  $q^2$  is divided into 1000 segments and in each segment the value of  $A_{\text{FB}}$  is computed. The search range for all the zero-crossing points discussed in this analysis was  $1.2 \text{ GeV}^2/c^4$ – $7.6 \text{ GeV}^2/c^4$ . The regions very close to the boundaries of the fit were omitted as the curve for  $A_{\text{FB}}$  has large uncertainties in these regions: the polynomials are less constrained due to their proximity to the start and end point. If the absolute value of  $A_{\text{FB}}$  is smaller than a given tolerance and the curve has a positive slope in this segment, this point is denoted a zero-crossing point and is kept. If there is more than one point fulfilling this criterion, the one closest to a reference value is kept. The reference value is chosen higher than the nominal value in both simulated samples, so that in practice the zero-crossing point with a higher  $q^2$  value is kept. In the collision data sample the situation of two zero-crossing points did not occur.

In about 99% of the cases a zero-crossing point with a positive slope can be determined on Standard Model samples of simulated events after reconstruction and with added background. About 1% of the events have two zero-crossing points with a positive slope. If all

<sup>7</sup>The described method is rather slow compared to numerically more elegant ones. However it is very robust and the time spent for determining the zero-crossing point compared to the execution time of the full algorithm is small.

zero-crossing points are considered (irrespective if they have a positive or negative slope), about 71% of the events have exactly one zero-crossing point. Furthermore, situations with no zero-crossing points exist as well. Examples of the different cases are given in Fig. 8.7.

To check that the quality of the fit is the same for events with one or multiple zero-crossing points and events with no zero-crossing point, the  $\chi^2$  value of the fits was determined, when the  $q^2$  range was split in eight bins. The results are displayed in Table 8.1 and are in very good agreement with each other. Note that a  $\chi^2$ -test is not an ideal method to check the absolute goodness-of-fit in these situations<sup>8</sup>. However, as only the values for the two cases are compared with respect to each other and the absolute values are not considered, this fact has no effect.

Table 8.1.:  $\chi^2$  values for fits to the forward and backward distributions, for the case where one or more zero-crossing point is found and for the case with no zero-crossing point. The degrees of freedom are the same for all numbers. No discrepancy in the fit quality can be observed.

Event type	$\chi^2$ value, forward events	$\chi^2$ value, backward events
one or more zero-crossing point	96.8	97.0
no zero-crossing point	94.1	97.4

## 8.5. Determination of the zero-crossing point in collision data

### 8.5.1. The numerical result

The zero-crossing point in collision data was determined using the technique described in the preceding section. Additionally, the fit is performed using the weights from the acceptance correction and the treatment of multiple candidates. The distributions of data points and the fitted PDFs are shown in Fig. 8.8. The  $A_{\text{FB}}$  curve, determined with the unbinned counting technique, in the range  $1.0 \text{ GeV}^2/c^4$  and  $7.8 \text{ GeV}^2/c^4$  is shown in Fig. 8.9. Additionally the result of a simple counting experiment for «forward» and «backward» is also shown. From this  $A_{\text{FB}}$  distribution, the zero-crossing point is determined to be:

$$q_0^2 = (4.9 \pm 0.9) \text{ GeV}^2/c^4 \quad (8.4)$$

<sup>8</sup>See Sect. 8.7 for more details on this issue.

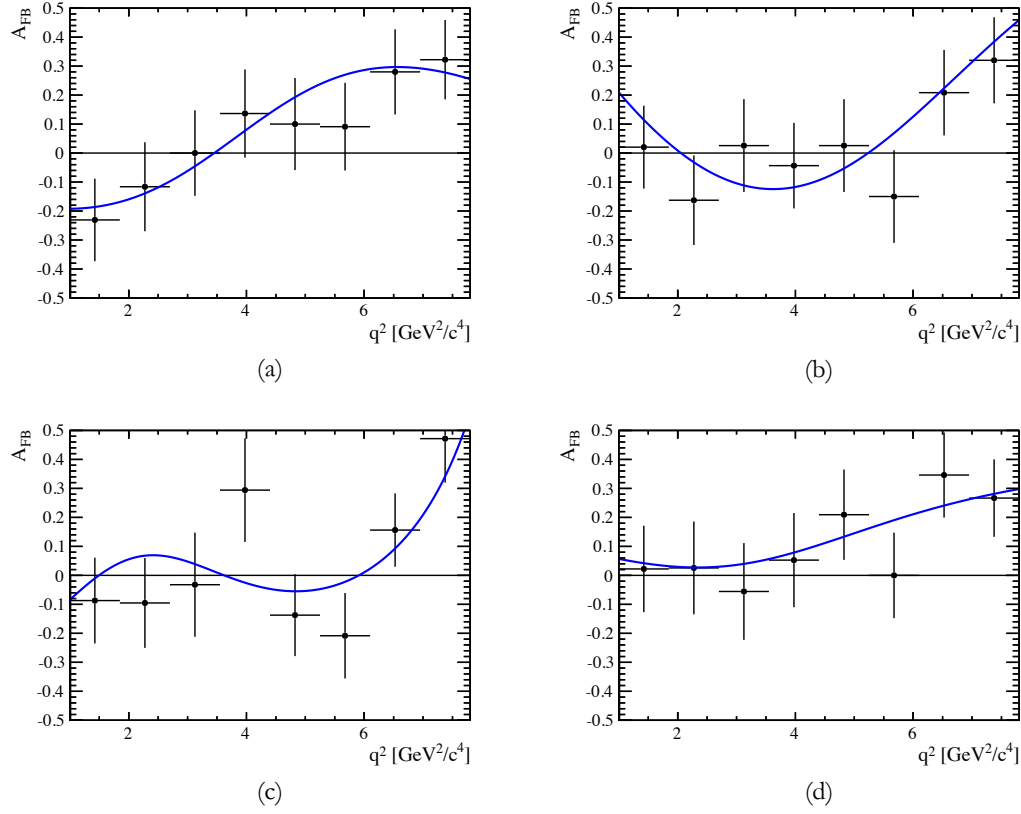


Figure 8.7.: Different possible situations for zero-crossing points in  $A_{\text{FB}}$  in simulation. The blue line is  $A_{\text{FB}}$  as determined with the unbinned counting technique, the black points are the binned  $A_{\text{FB}}$  of the signal only. With the definition used for the analysis on collision data, cases (a) to (c) would be considered having a valid zero-crossing point. In case (b) the second zero-crossing point would be taken (positive slope) and in case (c) the third one would be taken (positive slope with the largest value).

The (statistical) uncertainty is determined with a «bootstrapping» technique, which is described in the next section. The systematic uncertainty is negligible and will be discussed in Sect. 8.6.

This result can be compared with theoretical predictions for the zero-crossing point which have a central value in the range of  $4.0 - 4.3 \text{ GeV}^2/c^4$  [119] [121] [122]. Therefore the experimental result is in agreement with the theoretically predicted values. Furthermore the existence of a zero-crossing point rules out models which predicted a sign flip in the seventh Wilson coefficient,  $\mathcal{C}_7 = -\mathcal{C}_7^{SM}$ , see Sect. 5.5.

### 8.5.2. Bootstrapping

The statistical uncertainty was estimated using a «bootstrapping» technique [135]. The idea is the following: Given a (data) set  $\Omega$  with  $N$  events, one can construct sets  $\Omega_1, \dots, \Omega_m$  with  $N_1, \dots, N_m$  events, respectively. The events in  $\Omega_i$  are drawn randomly from  $\Omega$ , where each event can be drawn multiple times.  $N_1, \dots, N_m$  are distributed according to a Poisson distribution with mean value  $N$ .

For the estimation of the uncertainty of the zero-crossing point, the sample for forward- and backward events is used to construct sets  $\Omega_{F,1}, \dots, \Omega_{F,m}$  and  $\Omega_{B,1}, \dots, \Omega_{B,m}$ , where  $\Omega_{F,i}$  and  $\Omega_{B,i}$  stand for forward- and backward sample, respectively. Each sample is fitted with the PDF to determine the forward and backward distribution. For every pair of forward- and backward samples, the zero-crossing point is determined. This results in a distribution of zero-crossing points as a function of  $q^2$  which is shown in Fig. 8.10.

To determine the 68% confidence level, the range between  $0 \text{ GeV}^2/c^4$  and  $10 \text{ GeV}^2/c^4$  was divided in 100 bins. The bin with the largest number of entries was selected, and its contribution to the total integral covered by all bins was calculated. Then the bin with the second largest number of entries was added and the combined contribution to the total integral was calculated. This procedure was repeated until the covered integral was larger than 68% of the total integral. Note that this does not necessarily lead to a confidence interval which is connected: for certain distributions, the total interval is split up in several disjoint regions. However, for the distribution on collision data this was not the case. The confidence interval for the zero-crossing point, extracted with bootstrapping, ranges from  $4.0 \text{ GeV}^2/c^4$  to  $5.9 \text{ GeV}^2/c^4$ .

### 8.5.3. Fit-and-Toy technique

To cross-check the estimation of the uncertainty obtained with bootstrapping, a slightly different approach was performed as well. The PDFs for the forward and backward distributions were used as an input to a toy simulation. In this simulation, many data sets were created, where the events were distributed following the input PDFs and the number of events in the data sets were fluctuated following a Poisson distribution around the value measured in collision data. For all these samples the zero-crossing point was determined and the 68% confidence interval evaluated in the same way as for the bootstrapping. The

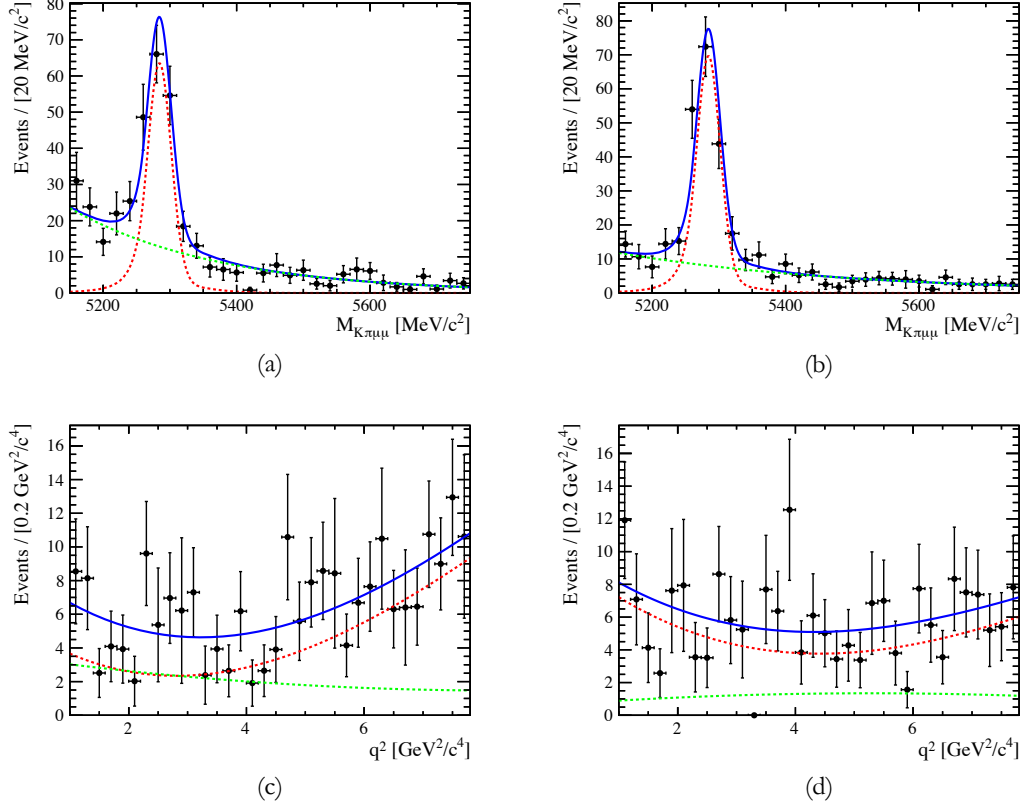


Figure 8.8.: Distributions of  $K\pi\mu\mu$  candidates masses for the forward (a) and backward (b) samples. Distribution of dimuon invariant masses squared,  $q^2$ , for forward (c) and backward (d) samples in collision data. The dashed red line shows the signal component and the dashed green line the background component. The blue line is the sum of both components. The figures for  $q^2$  integrate only over the signal region of the candidate masses and not over the full candidate mass region.

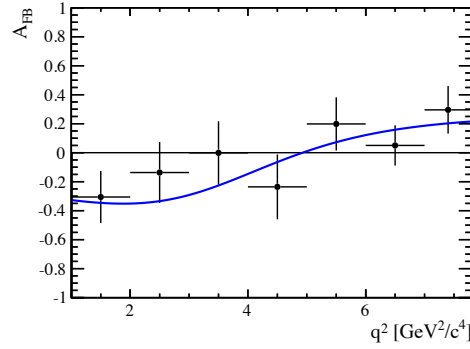
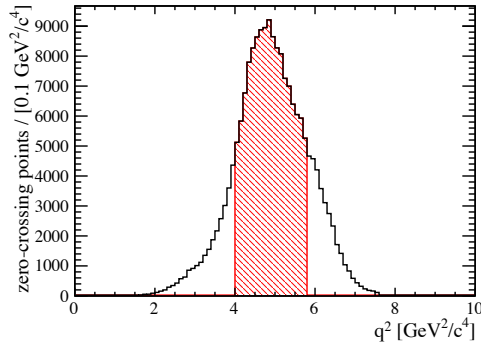
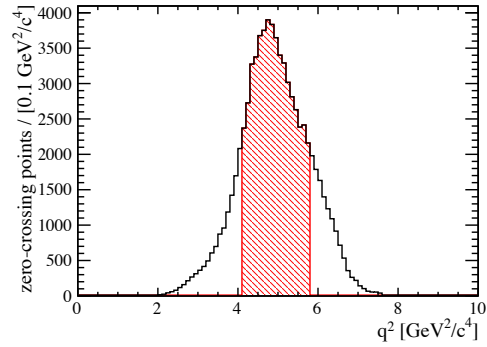


Figure 8.9.: Distribution of  $A_{FB}$  (blue line) from  $1 \text{ GeV}^2/c^4$  to  $7.8 \text{ GeV}^2/c^4$ . The black data points show the result of a counting experiment in bins of  $1 \text{ GeV}^2/c^4$ . The zero-crossing point determined from this distribution is  $4.9 \text{ GeV}^2/c^4$ .



(a)



(b)

Figure 8.10.: (a) shows the distribution of zero-crossing points as determined with a bootstrapping technique. The 68% confidence region is shown in red, the interval expands from  $4.0 \text{ GeV}^2/c^4$  to  $5.8 \text{ GeV}^2/c^4$ . (b) shows the distribution of zero-crossing points as determined with the fit-and-toy technique. The 68% confidence region is shown in red, the interval expands from  $4.1 \text{ GeV}^2/c^4$  to  $5.8 \text{ GeV}^2/c^4$ .

resulting interval ranges from  $4.1 \text{ GeV}^2/c^4$  to  $5.9 \text{ GeV}^2/c^4$  and is a bit more narrow than the one obtained with the bootstrapping but still in good agreement. The difference may be a consequence of randomising the weights in the bootstrapping which is not the case for this technique. The plot can be seen in Fig. 8.10.

## 8.6. Systematic uncertainties

The following sources of systematic uncertainties were considered:

1. Uncertainty in the IP smearing: The fit is redone with the IP smearing switched off.
2. Uncertainty in the binning of the PID variables: To account for this uncertainty, 50% of the events in the lowest 30% of a certain bin were migrated to the lower bin and 50% of the events in the highest 30% of the bin were migrated to the higher bin.
3. Uncertainty on the tracking efficiency: Possible systematic effects are taken into account by assigning the tracks with a momentum lower (higher) than  $10 \text{ GeV}/c$  an efficiency which is lower (higher) by one standard deviation and by assigning the tracks with a momentum higher than  $10 \text{ GeV}/c$  an efficiency which is higher (lower) by one standard deviation.
4. Uncertainty in the trigger efficiency: Systematic effects were accounted for by increasing or decreasing the trigger efficiency for muons with a momentum below  $3 \text{ GeV}/c$  by 3% for the acceptance correction.
5. Uncertainty of the isMuon criterium: The systematic uncertainty is assessed by fluctuating downwards the efficiency for tracks with a momentum less than  $10 \text{ GeV}/c$  by the statistical uncertainty and by fluctuating upwards the efficiency for tracks with a momentum more than  $10 \text{ GeV}/c$  by the statistical uncertainty. The procedure is also repeated by changing the direction of fluctuation for the corresponding two categories.
6. Acceptance correction: The acceptance correction was fluctuated by  $\pm 10\%$  for  $\cos \theta_\ell$ ,  $\cos \theta_K$  and both  $\cos \theta_\ell$ ,  $\cos \theta_K$  together. Furthermore the situation where the acceptance correction in  $\cos \theta_\ell$  and  $\cos \theta_K$  cannot be factorised is considered as well: The weights as a function of both angles are fluctuated upwards or downwards by  $\pm 10\%$  to account for this systematic effect.
7. The widths ( $\sigma$ ) of the Gaussian component of both Crystal Ball functions show a slight dependence on  $q^2$  which amounts to a slope corresponding to about 5%. These widths are varied by  $\pm 5\%$  in the fit and the result is recalculated.

Furthermore, some crosschecks were performed as well:

8. The fit was performed with and without reweighting the momentum of the  $B$  in the simulation to the values of the collision data.
9. The fit was performed with and without reweighting the transverse momentum of the  $B$  in the simulation to the values of the collision data.
10. The fit was performed with and without cutting on the momentum of  $3 \text{ GeV}/c$  on the hadrons.

The zero crossing points, evaluated under the changes to the data sample corresponding to the systematic checks, are listed in Table 8.2. Even when summing all systematic uncertainties and the deviations from the crosschecks in quadrature, which is a clear overestimation of the uncertainty, the overall systematic uncertainty is small compared to the statistical uncertainty and was not included in the overall uncertainty.

## 8.7. Goodness-of-Fit tests

To check the goodness of fit, the «point-to-point dissimilarity» method [136] [137] was used to calculate a p-value for the forward and the backward distribution. The result is:

$$p_{\text{forward}} = 61\% \quad (8.5)$$

$$p_{\text{backward}} = 73\% \quad (8.6)$$

which agrees well with the assumption that the PDF is compatible with the data points. More details on this method, its performance and also an application to the order of the polynomial which is used for the signal parametrisation in  $q^2$  can be found in Ref. [127].

## 8.8. Final result

The final result of the zero-crossing point determination of the forward-backward asymmetry with its uncertainty can be seen in Fig. 8.11. The numerical result, including statistical and systematic uncertainties, is  $q_0^2 = (4.9 \pm 0.9) \text{ GeV}^2/c^4$ . It is in good agreement with the theoretically predicted values in the range of  $4.0 - 4.3 \text{ GeV}^2/c^4$  [119] [121] [122].



Table 8.2.: Values for the zero-crossing point and deviation from the nominal value for all evaluations of the systematic uncertainty and the performed crosschecks. The type corresponds to the type given in the list of systematic uncertainties and crosschecks. The overall systematic uncertainty is calculated by adding all contributions (also the ones from the cross-checks) in quadrature. This is overly conservative, as shifts cannot occur in both directions at the same time. However, even with this procedure the systematic uncertainty is negligible with respect to the statistical.

Type	Zero-crossing point value [ $\text{GeV}^2/c^4$ ]	Deviation [ $\text{GeV}^2/c^4$ ]
1	4.92	0.01
2	4.93	0.00
	4.94	0.01
3	4.92	0.01
	4.93	0.00
4	4.93	0.00
	4.93	0.00
5	4.95	0.02
	4.92	0.01
	4.93	0.00
	4.92	0.01
	4.92	0.01
6	4.93	0.00
	4.92	0.01
	4.93	0.00
	4.95	0.02
	4.90	0.03
7	4.93	0.00
	4.92	0.01
8	4.94	0.01
9	4.93	0.00
10	4.94	0.01
total systematic uncertainty	—	0.05
total statistical uncertainty	—	0.9

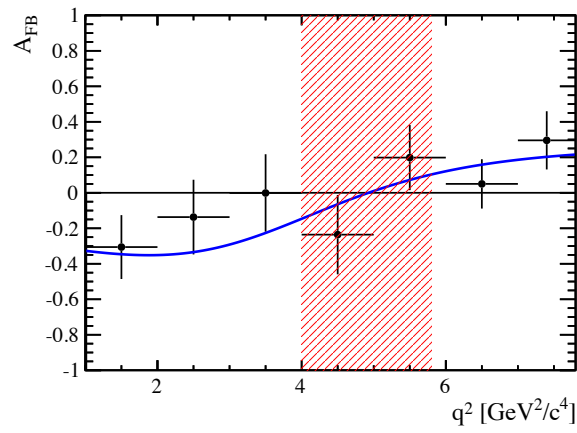


Figure 8.11.: Final result of the zero-crossing point determination of  $A_{FB}$ . In blue is  $A_{FB}$  as determined with the unbinned counting technique, the black data points correspond to the counting experiment and the red area is the 68% confidence level of statistical uncertainty on the zero-crossing point.

## 9. Measuring the $S_i$ and $P'_i$ observables in $B^0 \rightarrow K^{*0} \mu^+ \mu^-$

In the previous chapters, the understanding and preparation of the 2011 collision data sample was discussed: the acceptance correction was described and the comparison between simulated data and collision data was examined. This finally lead to the measurement of  $A_{FB}$ ,  $F_L$ ,  $A_T^{(2)}$ ,  $A_T^{Re}$ ,  $S_3$ ,  $S_9$ ,  $A_9$  and the zero-crossing point of  $A_{FB}$  [126]. The acquired understanding can now be exploited to measure the remaining (independent) observables of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . The task consists of studying the maximum likelihood fit to the collision data and the treatment of the uncertainties of the fit. Furthermore systematic effects have to be understood and their impact on the result has to be estimated.

This chapter presents the measurements of the observables  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  in the collision data collected in 2011. The  $P'_i$  observables and their angular distributions were introduced in Chap. 5. These observables are free of form-factor uncertainties in the low  $q^2$  region. The measurement was done with the same binning-scheme as the other angular analysis in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . The upper and lower boundaries of the six bins are listed in Table 9.1.

The measurement of the four  $P'_i$  observables is the first one done in any experiment.

### 9.1. Cross-checks with toy studies

In order to understand how well the  $P'_i$  observables can be extracted from the 2011 collision data set, several toy studies were performed. The angular PDFs for the  $S_i$  and  $P'_i$  observables for the angular fits and the generation of toy samples are described in Sect. 5.2 and 5.3.

#### 9.1.1. Pulls

To understand the behaviour of the fit and its uncertainty, unweighted toy samples were generated for 11  $P'_i$  values evenly distributed over the allowed range of the  $P'_i$  for each of the six bins in  $q^2$ . A minimal distance of 0.1 was kept to the boundaries in order not to generate too many toy samples with  $P'_i$  in the non-allowed region<sup>1</sup>. The values for  $F_L$  and  $A_T^{(2)}$  were set to the values obtained in the angular analysis of Ref. [126]. The results

<sup>1</sup>The non-allowed region is the region in phase-space where the PDF has a value below zero.

for the mean value of the fits and the  $\text{RMS}^2$  for  $P'_4$  are shown in Fig. 9.1. The RMS was preferred to the sigma of a Gaussian fit to the distribution as the distributions tended to be slightly asymmetric. Therefore a Gaussian fit would lead to a bias. Almost all central values and RMS agree within  $\pm 10\%$  with the input value. A larger deviation occurs at the boundaries of the second  $q^2$  bin where the difference amounts to approximately  $\pm 20\%$  of the uncertainty. However, as the value measured in collision data is not at the boundaries in this particular bin, this fact is not of importance.

The calculation of the pulls only includes the signal and no contamination due to background. Furthermore no weights were used in these samples as is the case in the data samples. Therefore, these pulls only serve as a guideline and not as an exact prediction for the bias of the central value and its uncertainty.

**A note on the boundaries** The boundaries for the pulls were calculated by sampling points in  $\theta_\ell$ ,  $\theta_K$  and  $\phi$  in a particular bin of  $q^2$ . For every sample, the PDF was evaluated and it was checked if it becomes negative. The region was considered unphysical if the PDF becomes negative at least once for all samples. This can only be a rough estimate, as it is possible to have a certain constellation of data points from collision data such that the PDF is positive even though it is in an unphysical region according to the sampling. This is the case for one value of  $P'_5$ , see later.

Table 9.1.: Binning scheme used for the angular analyses in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . The two gaps are around the  $q^2$  values of the  $J/\psi$  and  $\psi(2S)$  resonances. The first bin only starts at  $0.1 \text{ GeV}^2/c^4$  to avoid a rapidly changing acceptance correction. The last bin already stops at  $19 \text{ GeV}^2/c^4$  as the efficiency varies rapidly and there is not enough precision on the acceptance correction. The overall binning scheme is the same as used by the Belle collaboration.

Bin	Range [ $\text{GeV}^2/c^4$ ]
1	0.1 - 2.0
2	2.0 - 4.3
3	4.3 - 8.68
4	10.09 - 12.9
5	14.18 - 16.0
6	16.0 - 19.0

---

<sup>2</sup>RMS stands for «root mean square» and is defined as:  $\text{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)}$  where the  $x_i$  are the values in the set.

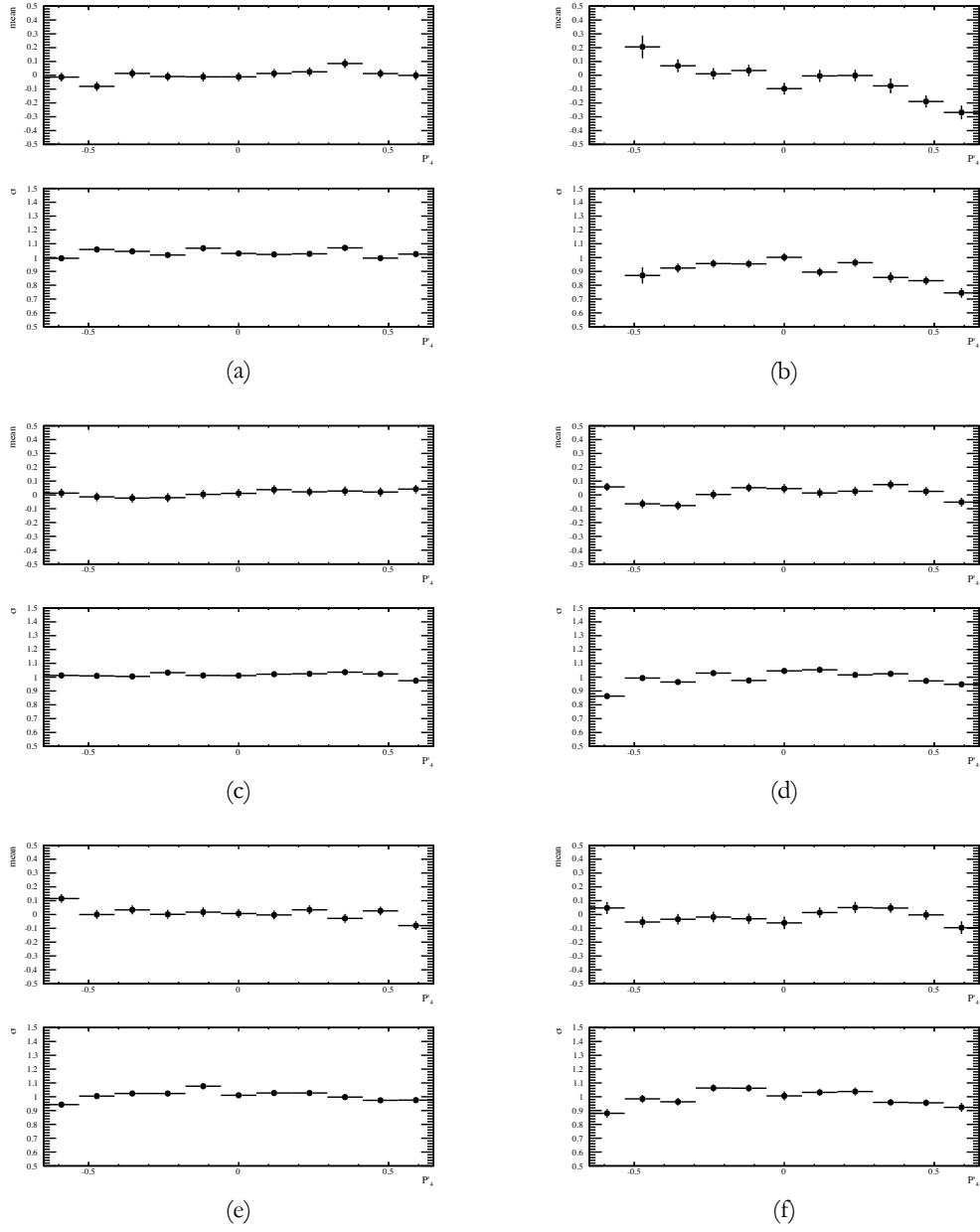


Figure 9.1.: Mean and RMS values for the pulls from toy samples of  $P'_4$  in all six  $q^2$  bins for 11 possible  $P'_4$  values in each bin. Except for the regions close to the boundary in bin 2 (b), the deviation is always smaller than 10%.

### 9.1.2. Values for $F_L$ and $A_T^{(2)}$

Although the values for  $F_L$  and  $A_T^{(2)}$  are not of primary interest in this measurement, they come as a by-product. Two cross-checks were performed. The first one was to calculate the pulls of  $F_L$  and  $A_T^{(2)}$  as it was done for the  $P'_i$  observables. The plots for  $P'_4$  are shown in Appendix F. There are two issues visible in these distributions which are shown in Fig. 9.2. The first one is for the second bin, where the mean of  $A_T^{(2)}$  is consistently too low in the full range of  $P'_4$ . The reason is that  $A_T^{(2)}$  appears multiplied with  $(1 - F_L)$  in the PDF. For large values of  $F_L$ ,  $(1 - F_L) \rightarrow 0$  and the sensitivity in  $A_T^{(2)}$  vanishes. This leads to a large spread of the values.

The second issue is present in the last  $q^2$  bin. As the value for  $A_T^{(2)}$  is low, some toy samples have a value of  $A_T^{(2)}$  very close to the mathematical boundary at  $-1$  which introduces an asymmetric distribution and therefore a bias in the mean.

As seen in the previous section, the pulls for the  $P'_4$  only show small deviations even though the pull values for  $A_T^{(2)}$  show a large bias. To understand the effect of this bias, the collision data was fitted where  $P'_i$ ,  $F_L$  and  $A_T^{(2)}$  were let floating and the value of  $P'_i$  was blinded<sup>3</sup>. In a second step, the values for  $F_L$  and  $A_T^{(2)}$  were set to the values fitted in the first iteration plus an offset composed of the relative deviation times the approximate uncertainty:

**Bin 2:**

$$A_T^{(2)} \rightarrow A_T^{(2)} + 0.5 \cdot 0.2 \quad (9.1)$$

$$F_L \rightarrow F_L - 0.07 \cdot 0.2 \quad (9.2)$$

**Bin 6:**

$$A_T^{(2)} \rightarrow A_T^{(2)} + 0.3 \cdot 0.6 \quad (9.3)$$

$$F_L \rightarrow F_L - 0.07 \cdot 0.1 \quad (9.4)$$

This leads to the differences shown in Table 9.2. The effect on  $P'_4$  is much smaller than the statistical precision which can be achieved with  $1 \text{ fb}^{-1}$  of collision data. The effect on the other  $P'_i$  observables is of a similar size.

### 9.1.3. Cross-check with counting experiments

While the  $P'_i$  observables cannot be obtained with a binned counting experiment, the  $S_i$  observables can. As measuring the  $S_i$  observables is experimentally almost identical to

---

<sup>3</sup>Blinding is a procedure to avoid accidental biases introduced by the experimenter by knowing the result in collision data and a desired result. The procedure is mostly employed in discovery searches to prevent selecting on fluctuations. The rule was less strict for this analysis as a bias could not easily be introduced with a fixed event selection. A blinding of the  $P'_i$  values was still put in place until the basic cross-checks were completed.

Table 9.2.: Results of a maximum likelihood fit on (blinded) collision data to estimate the impact of biases in  $A_T^{(2)}$  and  $F_L$  on the  $P'_i$  values. The first row for a bin is the outcome of the maximum likelihood fit with  $A_T^{(2)}$  and  $F_L$  left free, the second row for a bin is the outcome of the maximum likelihood fit when  $A_T^{(2)}$  and  $F_L$  were constrained according to the transformations in Eqs. 9.1 -9.4. The values themselves are blinded by an unknown offset so only the difference between the rows of a bin is relevant.

Type	$P'_4$	$F_L$	$A_T^{(2)}$
Bin 2, without transformation:	-0.85	0.83	-0.29
Bin 2, with transformation:	-0.84	0.816	-0.19
Bin 6, without transformation:	0.27	0.36	-0.67
Bin 6, with transformation:	0.26	0.353	-0.49

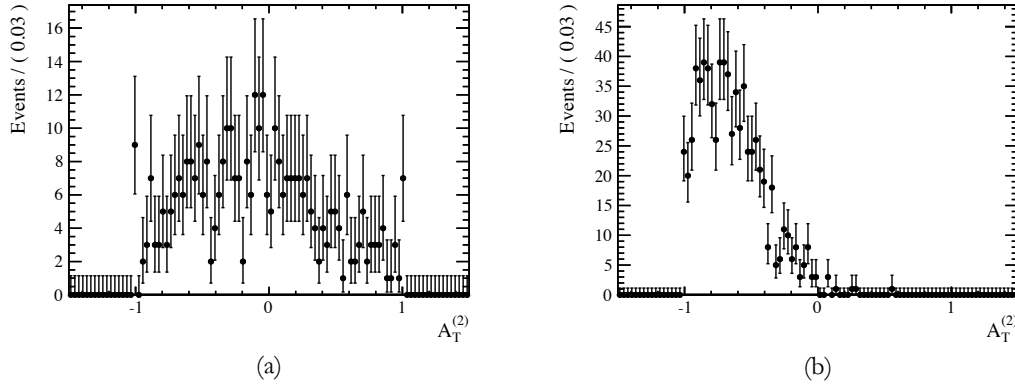


Figure 9.2.: Distributions of  $A_T^{(2)}$ , obtained with toy studies, in the second bin (a) and the sixth bin (b) of  $P'_4$ , illuminating the origin of the biases in the pull distribution of  $A_T^{(2)}$ . The boundaries of  $A_T^{(2)}$  are at 1.0 and -1.0.

measuring the  $P'_i$  observables, the counting experiment serves as an important cross-check for the whole analysis. To understand possible differences between a counting experiment and a unbinned maximum-likelihood fit to measure  $S_i$  a large toy sample was generated with the full angular distribution. The  $S_i$  were then determined by counting or by applying the corresponding folding and fitting. The counting experiment was performed by counting the number of «forward» and «backward» events, where forward and backward are defined according to the description in Sect. 5.2.5 and in Ref. [127]. An example of the obtained values with the nominal ones is given in Table 9.3. No large deviation can be seen and no bias present. This means that both methods reproduce the input value correctly.

Table 9.3.: Comparison of  $S_5$  values obtained with a counting experiment and with an angular fit, performed on a data sample generated with high statistics toy experiments. Both, the counting and the fitting, are in excellent agreement with the input value.

$S_5$ input value	$S_5$ counting	$S_5$ fitting
0.210	$0.209 \pm 0.001$	$0.210 \pm 0.001$
-0.130	$-0.131 \pm 0.001$	$-0.132 \pm 0.001$
-0.350	$-0.351 \pm 0.001$	$-0.352 \pm 0.001$
-0.380	$-0.380 \pm 0.001$	$-0.379 \pm 0.001$
-0.340	$-0.342 \pm 0.001$	$-0.340 \pm 0.001$
-0.250	$-0.250 \pm 0.001$	$-0.249 \pm 0.001$

## 9.2. Cross-checks with simulated data

Several cross-checks were performed with simulated data. One data set is obtained by simulating the decay of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  on «generator»-level. This means that the values obtained from this data set do not have resolution and reconstruction effects. The second one is data after «reconstruction»-level, where all the quantities are subject to detector resolution and acceptance. This data set represents the collision data with the same trigger conditions, stripping selection and BDT applied as in collision data. Both data sets are much larger than the  $1 \text{ fb}^{-1}$  of collision data collected in 2011. The idea of many corrections which are described in Chap. 6 is to get rid of experimental biases and measure the underlying physics distributions. The comparison between the results obtained on the two data sets therefore allows this assumption to be checked. The agreement between the results for the  $P'_i$  values, obtained on generator-level and on reconstruction-level in simulated data is shown in Fig. 9.3. As described in Sect. 8.3, the «folding» procedure was applied to



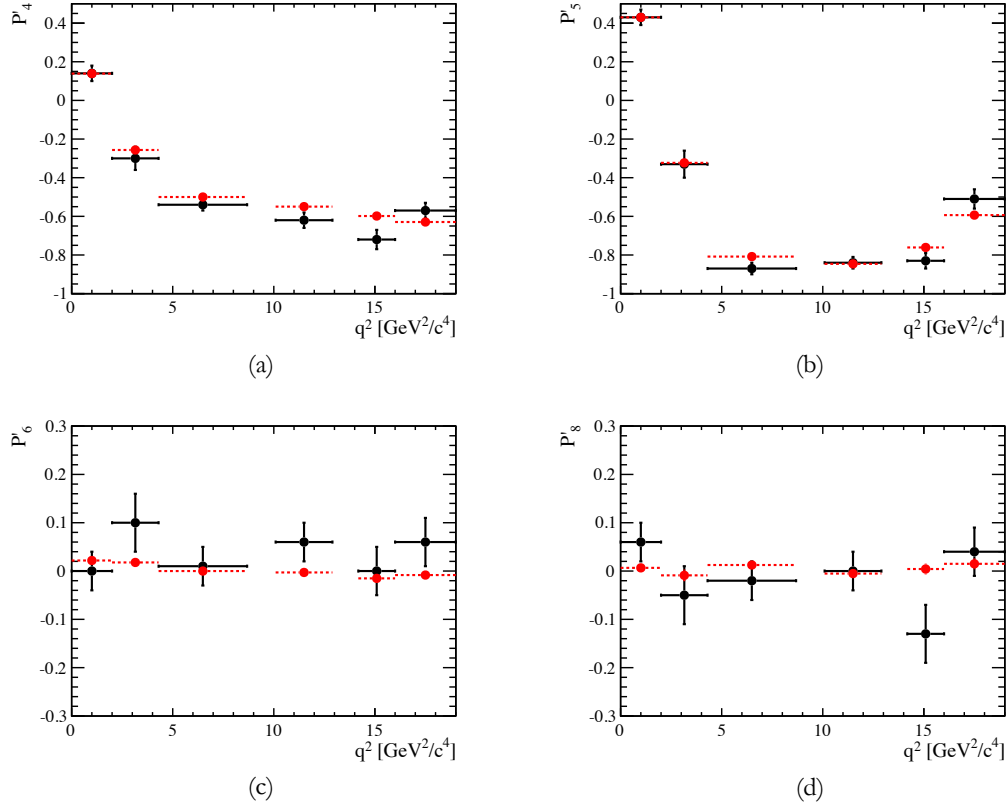


Figure 9.3.: Figures (a) to (d) show the comparison between an angular fit on reconstruction level (black data points) and generator-level (red data points). Overall there is good agreement, however note that the acceptance correction applied to the data sample on reconstruction level was not developed for such high statistics, so certain biases might be present.

the angles, leading to the simplified PDFs for the  $P'_i$ , and therefore must be applied to the data set as well.

### 9.2.1. Comparison between a fit and a counting experiment for the $S_i$ observables

The comparison between a counting experiment and a likelihood fit for the  $S_i$  was done on generator-level and on reconstruction-level simulation data. For the reconstruction-level fit in the individual  $q^2$  bins, the shape of the invariant mass was a double Crystal Ball function. The parameters for the double Crystal Ball function were taken from a fit to the invariant mass of the  $B^0$  candidates in the full sample of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . The results

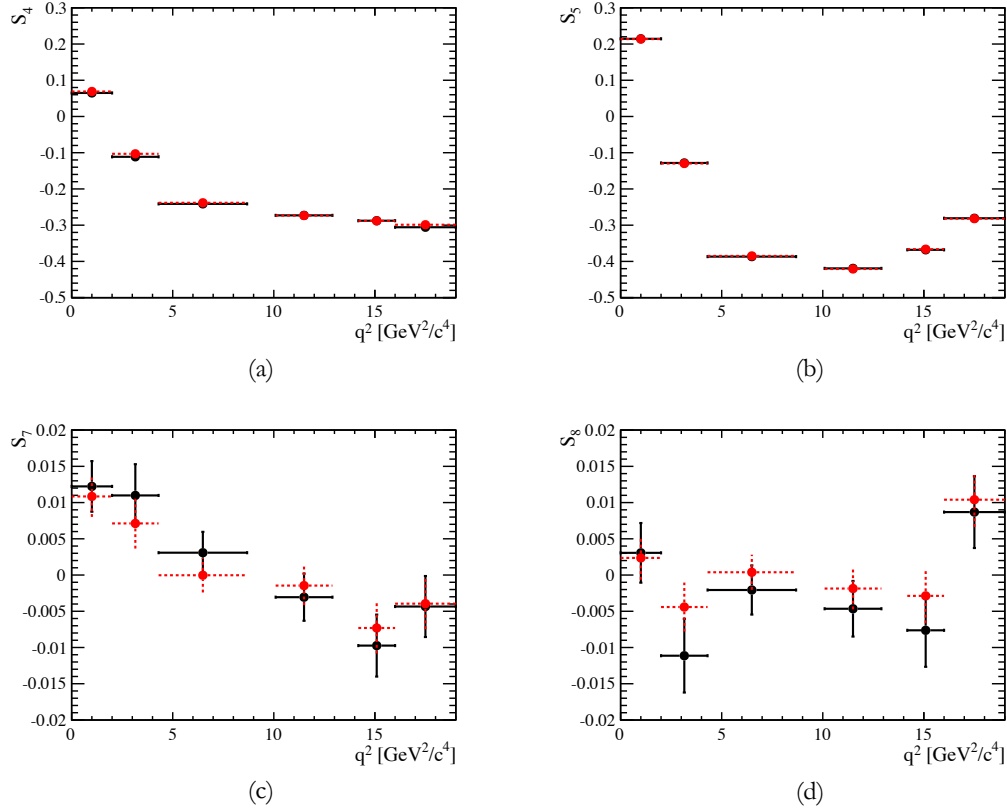


Figure 9.4.: Figure (a) to (d) show the comparison between a counting experiment (black data points) and an angular fit (red data points) for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  on a simulated sample on generator-level. There is good agreement for all four observables. The agreement for  $S_4$  and  $S_5$  is at the same level as the agreement for  $S_7$  and  $S_8$ .

for both studies are shown in Figs. 9.4 and 9.5. Both comparisons show a good agreement between the values from the counting experiment and the maximum likelihood fit.

### 9.2.2. Extracting the $S_i$ observables with and without folding

In order to cross-check if the folding introduces irregular effects not considered before, the sample on reconstruction level was fitted with the full PDF as described in Eq. 5.13, using Eqs. 5.18 and 5.19 and replacing the  $I_i$  terms with the  $S_i$  terms. The results for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  were then compared to the results obtained after folding the reconstruction-level data set. All numbers are in good agreement with each other and no systematic deviation could be found.

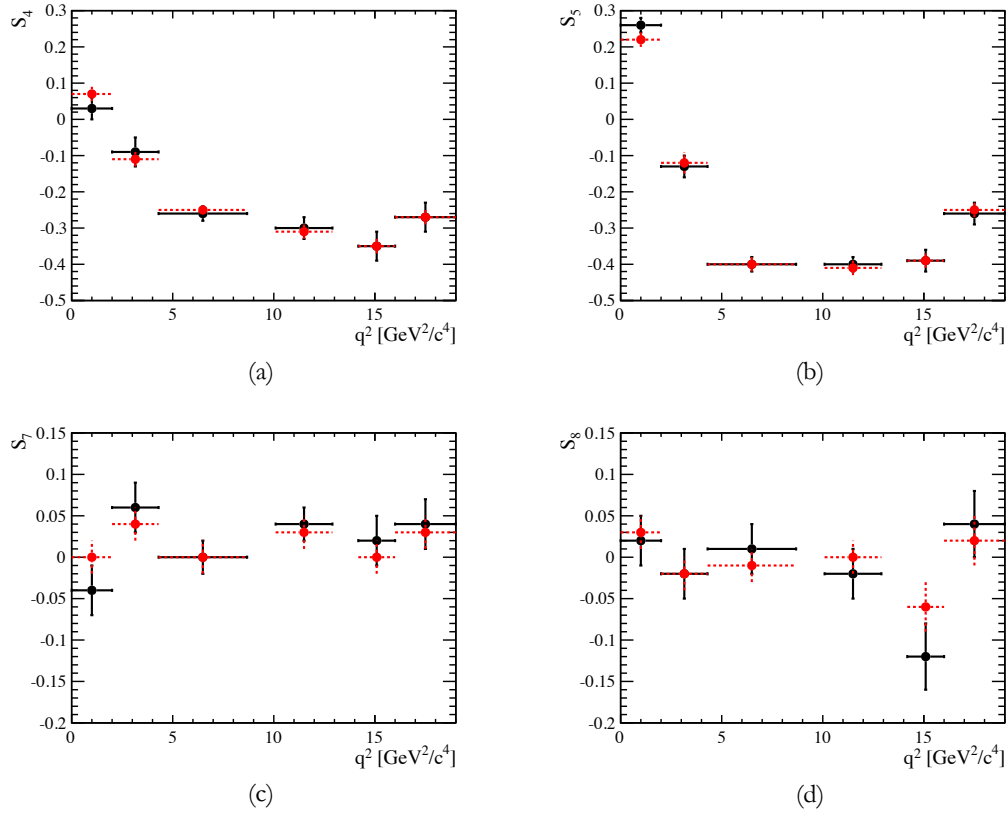


Figure 9.5.: Figure (a) to (d) show the comparison between a counting experiment (black data points) and an angular fit (red data points) for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  on a simulated sample on reconstruction-level. There is good agreement for all four observables. The agreement for  $S_4$  and  $S_5$  is at the same level as the agreement for  $S_7$  and  $S_8$ .

## 9.3. Fit on collision data

The sample of collision data was fitted in four dimensions, with one dimension being the invariant mass of the  $B^0$  candidate in order to distinguish signal from background candidates. The other three dimensions are  $\phi$ ,  $\cos \theta_\ell$  and  $\cos \theta_K$ . The invariant mass is fitted with a double Crystal Ball function for the signal component and an exponential function for the background component. The same shape was used as for the extraction of the zero-crossing point of  $A_{\text{FB}}$ , see Sect. 8.3. For the angles, the (folded) PDFs as described in Sect. 5.2.4, without an S-wave component, were used. For the background, second order polynomials were used in each angular variable. A non-extended maximum likelihood fit is used. As for the simulated samples, the data set had to be folded following the procedure as described for the PDFs in Sect. 8.3.

### 9.3.1. Calculation of the statistical uncertainty

The collision data is weighted to take different corrections into account, see Sect. 6.2. To calculate the statistical uncertainty on the fitted values, the profiled log likelihood is weighted with a scale factor  $\alpha$  which is:

$$\alpha = \frac{N}{\sum_{i=0}^N w_i}, \quad (9.5)$$

with  $N$  the number of events in the sample the fit is performed to and  $w_i$  the weight for event number  $i$ . The fit yields a value  $sPLL_{\text{best}}$  which corresponds to the minimal value of the scaled profile likelihood function. The upper and lower bounds on the 68% confidence level are found by scanning the scaled profile likelihood function until the deviation from  $sPLL_{\text{best}} + 0.5$  is the smallest. Two examples of scaled profile log likelihood functions are shown in Fig. 9.6. This procedure fails if the fitted value of the  $P'_i$  observable is close to a mathematical boundary with a value of the PDF being below zero. In this case the  $sPLL$  cannot be evaluated anymore. An alternative approach is performed in these situations. Instead of scanning the scaled profile log likelihood, it is assumed to be symmetric. The uncertainty can then be approximated by calculating the covariance matrix  $V'$ :

$$V' = VC^{-1}V \quad (9.6)$$

where  $V$  is the covariance matrix calculated from a fit to  $-\log L$  and  $C$  is the covariance matrix calculated from a fit to  $-\log L'$ . The expressions  $-\log L$  and  $-\log L'$  are defined as:

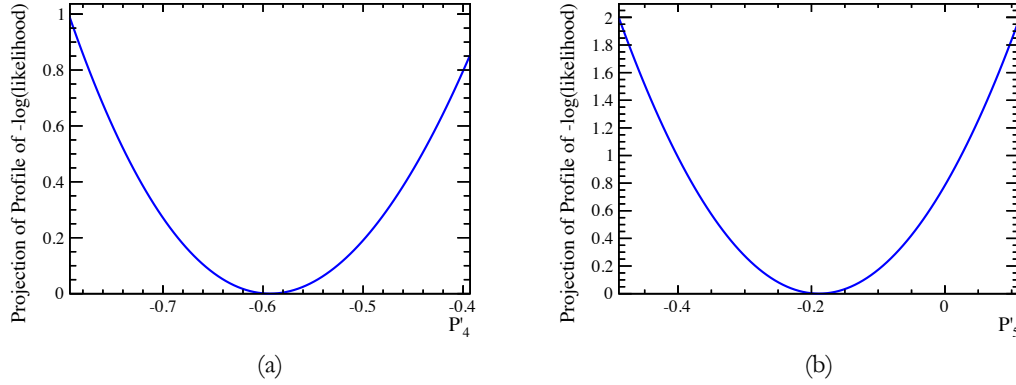


Figure 9.6.: Figure (a) shows the projection of the profile likelihood of  $P'_4$  in bin 3. Note the slightly asymmetric curve. Figure (b) shows the projection of the profile likelihood of  $P'_5$  in bin 3.

$$-\log L = -\sum_{i=1}^N w_i \log(\text{pdf}(x_i)) \quad (9.7)$$

$$-\log L' = -\sum_{i=1}^N w_i^2 \log(\text{pdf}(x_i)) \quad (9.8)$$

In situations where  $P'_i$  is far away from the boundary, both methods yield results which are very close.

### 9.3.2. Fit results

The results of the maximum likelihood fits for all four  $P'_i$  are shown in Table 9.4. All uncertainties are calculated by scanning the likelihood except the fourth bin for  $P'_5$  and the lower bound for the sixth bin of  $P'_5$  which were calculated using the covariance matrices of the fit. In these cases, it is possible that the negative error expands beyond the boundary. Furthermore, the sixth bin of  $P'_5$  is outside the allowed region when calculated with the method described in Sect. 9.1.1. However, the fit still converges normally. This is most likely due to a statistical fluctuation with the particular values realised in collision data being an «allowed island» in the non-allowed region. An example of a fit projection in the invariant mass and the three angles is given in Fig. 9.7.

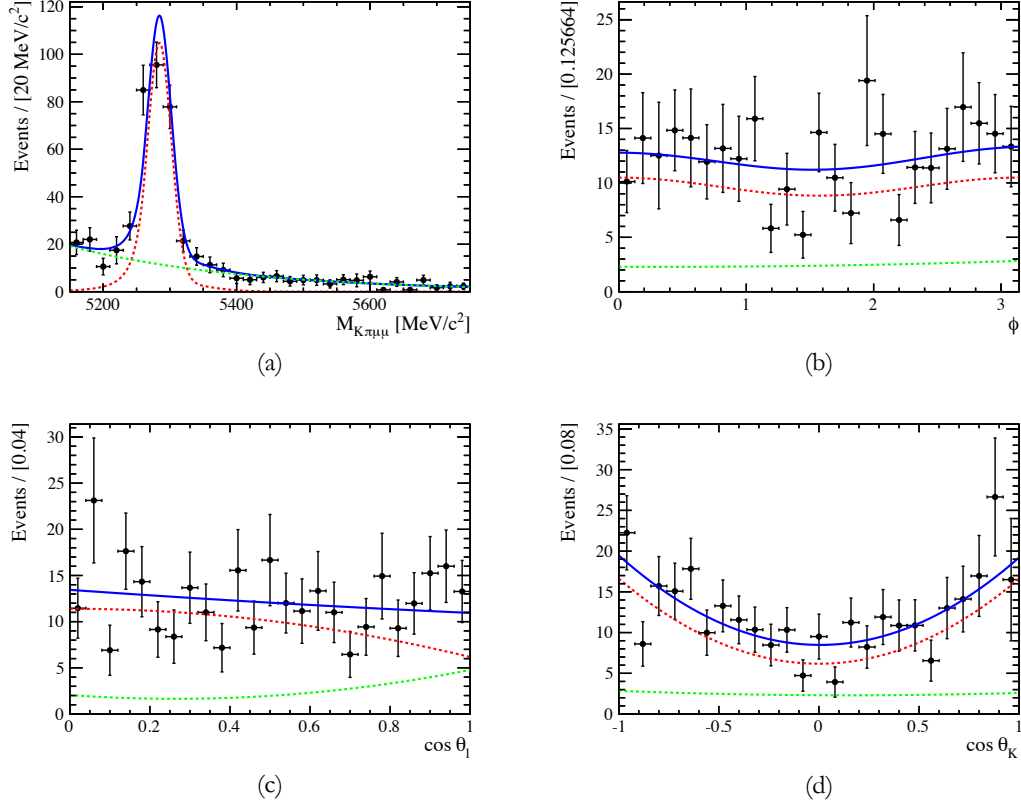


Figure 9.7.: The figure shows the results of the maximum likelihood fit in the third bin of  $P'_5$ . The blue line is the curve of the overall PDF, the green dashed line is the background component and red dashed line is the signal component. The black data points are the measured values from collision data. The invariant mass of the  $B^0$  candidates is shown in (a), the angular distributions in  $\phi$ ,  $\cos \theta_\ell$  and  $\cos \theta_K$  are shown in (b), (c) and (d) respectively. The figures for the angular distributions integrate only over the signal region of the candidate masses and not over the full candidate mass region.

Table 9.4.: Central values and statistical uncertainties corresponding to the 68% confidence region for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$ .

Bin	$P'_4$	$P'_5$	$P'_6$	$P'_8$
1	$0.00^{+0.20}_{-0.20}$	$0.45^{+0.18}_{-0.19}$	$-0.24^{+0.22}_{-0.21}$	$-0.06^{+0.24}_{-0.25}$
2	$-0.37^{+0.49}_{-0.49}$	$0.29^{+0.46}_{-0.41}$	$0.15^{+0.45}_{-0.42}$	$-0.15^{+0.48}_{-0.50}$
3	$-0.59^{+0.15}_{-0.14}$	$-0.19^{+0.15}_{-0.15}$	$-0.04^{+0.15}_{-0.15}$	$0.29^{+0.15}_{-0.15}$
4	$-0.46^{+0.21}_{-0.21}$	$-0.79^{+0.20}_{-0.20}$	$-0.31^{+0.22}_{-0.21}$	$-0.06^{+0.22}_{-0.22}$
5	$0.09^{+0.25}_{-0.25}$	$-0.79^{+0.21}_{-0.19}$	$-0.18^{+0.23}_{-0.22}$	$-0.20^{+0.21}_{-0.21}$
6	$-0.35^{+0.22}_{-0.21}$	$-0.60^{+0.18}_{-0.18}$	$0.31^{+0.22}_{-0.22}$	$0.06^{+0.24}_{-0.24}$

### 9.3.3. Comparison with counting experiment

As for the simulated data, the  $S_i$  observables can be compared when obtained with a maximum likelihood fit and with a counting experiment. The comparison is shown in Fig. 9.8. All four observables show a good agreement. This comparison serves as an important crosscheck for the results themselves, as counting is not affected by the parametrisations of the angular variables and by any difficulties due to the maximum likelihood fit.

## 9.4. Estimating the S-wave contribution

As described in Sect. 5.2.6 an underlying S-wave contribution could alter the angular distribution in the three angles and, therefore, the final result. As there was not enough data available in the individual bins in  $q^2$ , an upper limit for the S-wave contribution was estimated. For this purpose toy samples were generated where  $F_S$  was set to 7% which is the upper limit at 68% confidence level of a study done for the first angular analysis of  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ .  $A_S$  and  $A_S^{(i)}$  were set to their upper and lower limits as described in Eqs. 5.49 - 5.51. The number of events in each  $q^2$  was set to the value obtained in collision data. The toy samples were then fitted with the PDF for the  $P'_i$  without the S-wave contribution. The results of this study and the differences with the nominal (input) values are shown in Tables 9.5 - 9.8. For the systematic uncertainty, the largest deviation from the nominal value was considered and added quadratically to the lower or upper error of the systematic uncertainty, depending if the deviation due to the S-wave is positive or negative. Note that this is most likely an overestimation of the systematic uncertainty as the values for  $A_S$  and  $A_S^{(i)}$  might not be at their possible boundaries in collision data. However, in most of the bins the discrepancy is small compared to the overall statistical uncertainty.

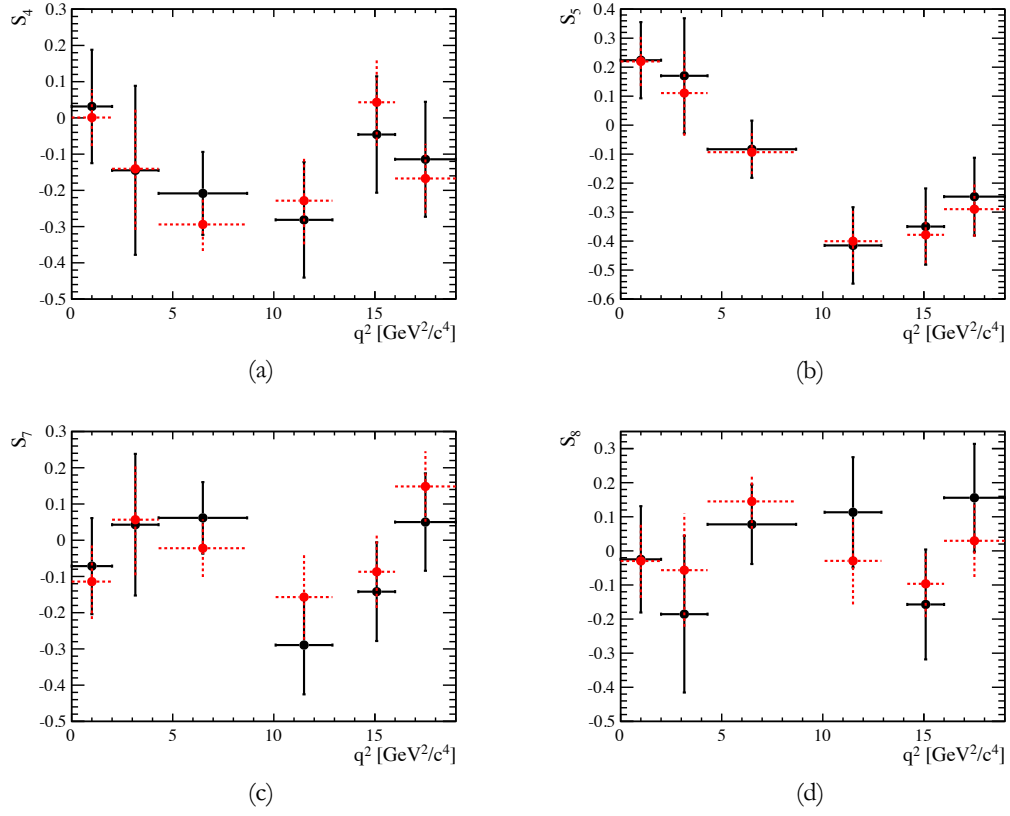


Figure 9.8.: Figure (a) to (d) show the comparison between a counting experiment (black data points) and an angular fit (red data points) for  $S_4$ ,  $S_5$ ,  $S_7$  and  $S_8$  on collision data. There is good agreement for all four observables.



Table 9.5.: Results for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  when generating toy samples with an S wave contribution and fitting without.  $A_S$  and the  $A_S^{(i)}$  were put to their respective upper boundaries. Note that for  $P'_8$  the term with  $A_S$  is not present.  $\Delta P'_i$  is the input value for  $P'_i$  with the S wave minus the fitted result without. In most of the cases, the fitted result underestimates the deviation from zero.

Bin	$P'_{4,fit}$	$\Delta P'_4$	$P'_{5,fit}$	$\Delta P'_5$	$P'_{6,fit}$	$\Delta P'_6$	$P'_{8,fit}$	$\Delta P'_8$
1	-0.032	0.002	0.417	0.033	-0.225	-0.035	-0.070	-0.010
2	-0.333	-0.037	0.286	0.004	0.148	0.012	-0.128	-0.012
3	-0.560	-0.040	-0.168	-0.022	-0.038	-0.012	0.254	0.016
4	-0.403	-0.027	-0.599	-0.191	-0.305	-0.045	-0.057	-0.013
5	0.099	0.001	-0.611	-0.179	-0.150	-0.020	-0.193	-0.017
6	-0.316	-0.024	-0.519	-0.081	0.308	0.012	0.053	-0.003

Table 9.6.: Results for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  when generating toy samples with an S wave contribution and fitting without.  $A_S$  was put to the lower boundary while the  $A_S^{(i)}$  were put to their respective upper boundaries. Note that for  $P'_8$  the term with  $A_S$  is not present.  $\Delta P'_i$  is the input value for  $P'_i$  with the S wave minus the fitted result without. In most of the cases, the fitted result underestimates the deviation from zero.

Bin	$P'_{4,fit}$	$\Delta P'_4$	$P'_{5,fit}$	$\Delta P'_5$	$P'_{6,fit}$	$\Delta P'_6$	$P'_{8,fit}$	$\Delta P'_8$
1	-0.029	-0.001	0.365	0.085	-0.244	-0.016	-0.070	-0.010
2	-0.348	-0.022	0.250	0.040	0.135	0.025	-0.128	-0.012
3	-0.569	-0.031	-0.177	-0.013	-0.056	0.006	0.254	0.016
4	-0.412	-0.018	-0.739	-0.051	-0.319	-0.031	-0.057	-0.013
5	0.081	0.019	-0.731	-0.059	-0.154	-0.016	-0.193	-0.017
6	-0.317	-0.023	-0.563	-0.037	0.222	0.098	0.053	-0.003

Table 9.7.: Results for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  when generating toy samples with an S wave contribution and fitting without.  $A_S$  was put to the upper boundary while the  $A_S^{(i)}$  were put to their respective lower boundaries. Note that for  $P'_8$  the term with  $A_S$  is not present.  $\Delta P'_i$  is the input value for  $P'_i$  with the S wave minus the fitted result without. In most of the cases, the fitted result underestimates the deviation from zero.

Bin	$P'_{4,fit}$	$\Delta P'_4$	$P'_{5,fit}$	$\Delta P'_5$	$P'_{6,fit}$	$\Delta P'_6$	$P'_{8,fit}$	$\Delta P'_8$
1	-0.047	0.017	0.374	0.067	-0.247	-0.013	-0.075	-0.005
2	-0.358	-0.012	0.242	0.048	0.141	0.019	-0.139	-0.001
3	-0.561	-0.039	-0.178	-0.012	-0.057	0.007	0.258	0.012
4	-0.412	-0.018	-0.742	-0.048	-0.324	-0.026	-0.057	-0.013
5	0.090	0.010	-0.736	-0.054	-0.151	-0.019	-0.206	0.004
6	-0.322	-0.018	-0.557	-0.043	0.227	0.093	0.050	0.000

Table 9.8.: Results for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  when generating toy samples with an S wave contribution and fitting without.  $A_S$  and the  $A_S^{(i)}$  were put to their respective lower boundaries. Note that for  $P'_8$  the term with  $A_S$  is not present.  $\Delta P'_i$  is the input value for  $P'_i$  with the S wave minus the fitted result without. In most of the cases, the fitted result underestimates the deviation from zero.

Bin	$P'_{4,fit}$	$\Delta P'_4$	$P'_{5,fit}$	$\Delta P'_5$	$P'_{6,fit}$	$\Delta P'_6$	$P'_{8,fit}$	$\Delta P'_8$
1	-0.026	-0.004	0.416	0.034	-0.223	-0.037	-0.075	-0.005
2	-0.351	-0.019	0.286	0.004	0.150	0.010	-0.139	-0.001
3	-0.548	-0.052	-0.180	-0.010	-0.035	-0.015	0.258	0.012
4	-0.400	-0.003	-0.601	-0.189	-0.315	-0.035	-0.057	-0.013
5	0.089	0.011	-0.603	-0.187	-0.146	-0.024	-0.206	-0.004
6	-0.351	0.011	-0.523	-0.077	0.316	0.004	0.050	0.000

## 9.5. Systematic uncertainties

The systematic uncertainties for the  $P'_i$  observables were evaluated in the same way as for the systematic uncertainty of the zero-crossing point of  $A_{\text{FB}}$ , see Sect. 8.6. The individual values, obtained by fitting the data samples corresponding to the systematic checks, are listed in Tables G.1, G.2, G.3 and G.4. For the calculation of the overall systematic uncertainty, a slightly different procedure compared to the determination of the zero-crossing point of  $A_{\text{FB}}$  was chosen<sup>4</sup>. The reasoning was the following: it is not possible that a systematic shift is in both directions, up and down. The idea was then to take the largest deviation from the fitted value for every category of systematics (as defined in Sect. 8.6). This means that from category 1 to 5, the largest deviation was taken; for category 6 the largest deviation from the first six values and from the last two values was taken; from categories 7 and 9 the largest deviation was taken. The deviations were then summed quadratically. The categories 8 and 10 were considered a cross-check which would only have been taken into account, if large deviations would have occurred. This was not the case. The overall systematic uncertainties resulting from this procedure are given in Table 9.9.

Table 9.9.: Overall systematics for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$ . The procedure to extract these values from the full systematic analysis is described in the text.

Bin	$P'_4$	$P'_5$	$P'_6$	$P'_8$
1	0.024	0.022	0.025	0.019
2	0.071	0.055	0.038	0.071
3	0.014	0.019	0.047	0.020
4	0.012	0.037	0.013	0.018
5	0.029	0.022	0.019	0.015
6	0.020	0.033	0.021	0.026

## 9.6. Experimental results and comparison with theoretical predictions

The experimental numbers, including statistical uncertainties, systematic uncertainties and uncertainties due to the S-wave contribution are listed in Table 9.10. Note that the uncertainty due to the S-wave contribution is only added in one direction as it leads to a unidirectional shift.

<sup>4</sup>The procedure for the zero-crossing point is overly conservative, but it does not matter as its effect on the result is negligible.

Table 9.10.: Experimental results for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$ . The first uncertainty is the statistical, the second uncertainty is the systematic. The third uncertainty is only one-sided and stems from the influence of the S-wave.

Bin	$P'_4$	$P'_5$	$P'_6$	$P'_8$
1	$0.00^{+0.20+0.02+0.02}_{-0.20-0.02}$	$0.45^{+0.18+0.02+0.09}_{-0.19-0.02}$	$-0.24^{+0.22+0.03}_{-0.21-0.03-0.04}$	$-0.06^{+0.24+0.02}_{-0.25-0.02-0.01}$
2	$-0.37^{+0.49+0.07}_{-0.49-0.07-0.04}$	$0.29^{+0.46+0.06+0.05}_{-0.41-0.06}$	$0.15^{+0.45+0.04+0.03}_{-0.42-0.04}$	$-0.15^{+0.48+0.07}_{-0.50-0.07-0.01}$
3	$-0.59^{+0.15+0.01}_{-0.14-0.01-0.05}$	$-0.19^{+0.15+0.02}_{-0.15-0.02-0.02}$	$-0.04^{+0.15+0.05}_{-0.15-0.05-0.02}$	$0.29^{+0.15+0.02+0.02}_{-0.15-0.02}$
4	$-0.46^{+0.21+0.01}_{-0.21-0.01-0.03}$	$-0.79^{+0.20+0.04}_{-0.20-0.04-0.19}$	$-0.31^{+0.22+0.01}_{-0.21-0.01-0.05}$	$-0.06^{+0.22+0.02}_{-0.22-0.02-0.01}$
5	$0.09^{+0.25+0.03+0.02}_{-0.25-0.03}$	$-0.79^{+0.21+0.02}_{-0.19-0.02-0.19}$	$-0.18^{+0.23+0.02}_{-0.22-0.02-0.02}$	$-0.20^{+0.21+0.02}_{-0.21-0.02-0.02}$
6	$-0.35^{+0.22+0.02}_{-0.21-0.02-0.02}$	$-0.60^{+0.18+0.03}_{-0.18-0.03-0.08}$	$0.31^{+0.22+0.02+0.10}_{-0.22-0.02}$	$0.06^{+0.24+0.03}_{-0.24-0.03-0.00}$

Table 9.11.: Theoretical predictions for  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  in the Standard Model, based on Ref. [138].

Bin	$P'_4$	$P'_5$	$P'_6$	$P'_8$
1	$0.171^{+0.010+0.009}_{-0.013-0.009}$	$0.533^{+0.028+0.017}_{-0.036-0.020}$	$0.084^{+0.035+0.026}_{-0.021-0.026}$	$0.019^{+0.013+0.013}_{-0.008-0.013}$
2	$-0.285^{+0.030+0.011}_{-0.035-0.010}$	$-0.334^{+0.095+0.020}_{-0.111-0.019}$	$0.098^{+0.046+0.031}_{-0.030-0.031}$	$0.035^{+0.019+0.012}_{-0.012-0.013}$
3	$-0.502^{+0.008+0.015}_{-0.007-0.012}$	$-0.872^{+0.043+0.030}_{-0.029-0.029}$	$0.027^{+0.021+0.059}_{-0.012-0.059}$	$0.010^{+0.011+0.025}_{-0.005-0.027}$
4	$-0.541^{+0.072+0.009}_{-0.070-0.007}$	$-0.893^{+0.223+0.018}_{-0.110-0.017}$	$-0.001^{+0.004+0.034}_{-0.003-0.034}$	$-0.008^{+0.005+0.014}_{-0.003-0.015}$
5	$-0.581^{+0.166+0.004}_{-0.095-0.008}$	$-0.779^{+0.328+0.010}_{-0.363-0.009}$	$0.0^{+0.0+0.0}_{-0.0-0.0}$	$-0.008^{+0.005+0.002}_{-0.006-0.002}$
6	$-0.632^{+0.124+0.002}_{-0.060-0.002}$	$-0.601^{+0.282+0.008}_{-0.367-0.007}$	$0.0^{+0.0+0.0}_{-0.0-0.0}$	$-0.004^{+0.003+0.001}_{-0.004-0.001}$

The theoretical predictions were taken from Ref. [138]. The comparison with the experimental values are shown in Fig. 9.9. While there is a good agreement for  $P'_6$  and  $P'_8$ , there is a slight discrepancy between the experimental numbers and their predictions of  $P'_4$  in the high  $q^2$  region. There is a large discrepancy for  $P'_5$  in the low  $q^2$  region. Several cross-checks to address a possible experimental origin of the discrepancy in  $P'_5$  are described in the next section.

## 9.7. Further cross-checks

### 9.7.1. Comparison between $S_i$ and $P'_i$ observables

As the values for  $P'_i$  can be calculated using the values for  $S_i$  and  $F_L$  which are obtained from an angular fit in a given bin, a consistency check was performed. Its result is listed

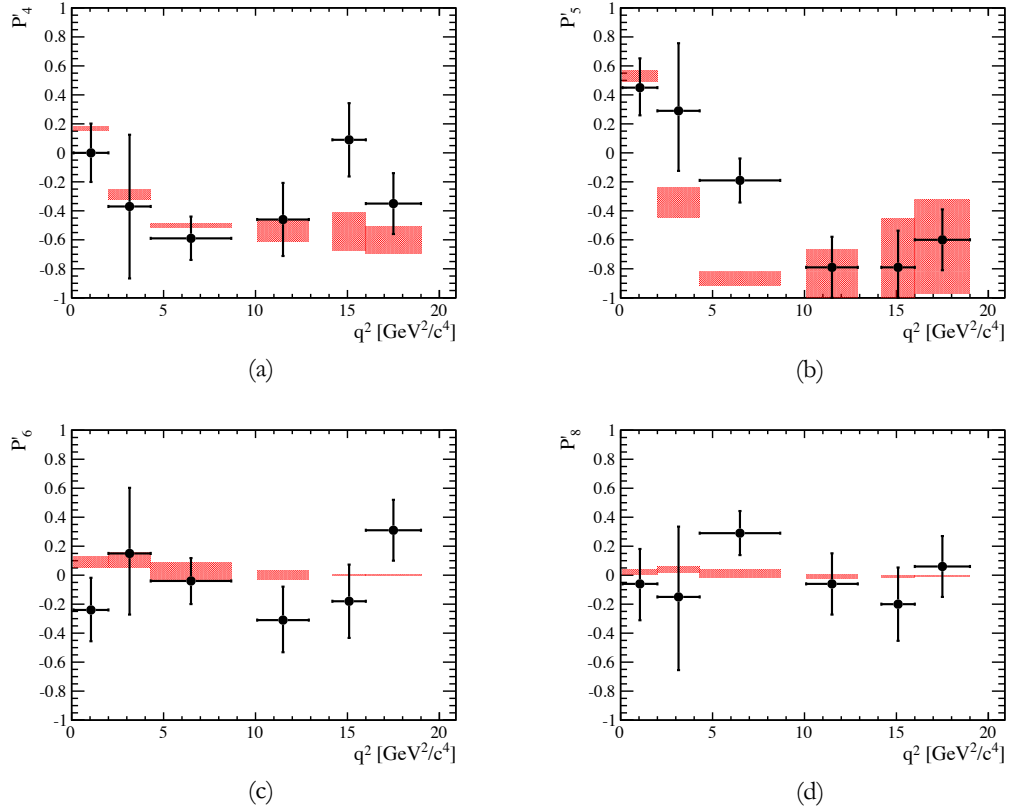


Figure 9.9.: Comparison of experimental values of  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$ , with statistical, systematic and S-wave uncertainties added quadratically, and theoretical values. The uncertainties for the theoretical values in the two last  $q^2$  bins of  $P'_6$  are negligible, the small error bar is drawn for visualisation purposes only.

in Table 9.12 and 9.13. There is a very good agreement between the values of the  $P'_i$  stemming from the fit and the calculated  $P'_i$  values from the fitted  $S_i$  and  $F_L$  values. The discrepancies are negligible compared to the experimental precision.

### 9.7.2. Further checks for the third bin of $P'_5$

The following checks were done to check that the tension in the third bin of  $P'_5$  is not due to wrong parametrisation of the data.

- The background was parametrised with a first order polynomial instead of a second order. The effect on the value of  $P'_5$  is negligible.
- The left sideband was not included in the fit *i.e.* the fit started at 5220 MeV/ $c^2$  in the mass. This had a negligible influence on the result.
- Fitting without weights: The data was fitted without using the weights. This changes the value of  $P'_5$  from  $-0.19$  to  $-0.27$ . However, not using the weights is a systematic mistake and a large overestimation of the uncertainty.
- The upper limit of the third bin was changed from 8.68 GeV $^2/c^4$  to 8.0 GeV $^2/c^4$ . This changes the value of  $P'_5$  from  $-0.19$  to  $-0.28$ . While this is certainly an effect, it is much too small to explain the discrepancy and could just as well be a statistical fluctuation.

Table 9.12.: Table with cross-checks for the fitted  $P'_4$  and  $P'_5$  values. The second and fourth columns show the values of  $P'_4$  and  $P'_5$  in the six  $q^2$  bins obtained with an angular fit with the PDF for  $P'_4$  and  $P'_5$ . The third and fifth columns show the derived values of  $P'_4$  and  $P'_5$  in the six  $q^2$  bins using an angular fit and deriving the values from  $F_L$ ,  $S_4$  and  $S_5$ . All values show a very good agreement.

Bin	Value of $P'_4$	Value of $\frac{S_4}{\sqrt{F_L(1-F_L)}}$	Value of $P'_5$	Value of $\frac{S_5}{\sqrt{F_L(1-F_L)}}$
1	0.00	0.00	0.45	0.45
2	-0.37	-0.37	0.29	0.29
3	-0.59	-0.59	-0.19	-0.19
4	-0.46	-0.46	-0.79	-0.80
5	0.09	0.09	0.79	0.79
6	-0.35	-0.35	-0.60	-0.60

Table 9.13.: Table with cross-checks for the fitted  $P'_6$  and  $P'_8$  values. The second and fourth columns show the values of  $P'_6$  and  $P'_8$  in the six  $q^2$  bins obtained with an angular fit with the PDF for  $P'_6$  and  $P'_8$ . The third and fifth columns show the derived values of  $P'_6$  and  $P'_8$  in the six  $q^2$  bins using an angular fit and deriving the values from  $F_L$ ,  $S_7$  and  $S_8$ . All values show a very good agreement.

Bin	Value of $P'_6$	Value of $\frac{S_7}{\sqrt{F_L(1-F_L)}}$	Value of $P'_8$	Value of $\frac{S_8}{\sqrt{F_L(1-F_L)}}$
1	-0.24	-0.24	-0.06	-0.06
2	0.15	0.15	-0.15	-0.14
3	-0.04	-0.04	0.29	0.29
4	-0.31	-0.31	-0.06	-0.06
5	-0.18	-0.18	-0.20	-0.20
6	0.31	0.31	0.06	0.06

## 9.8. Conclusion

The observables  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  have been measured for the first time experimentally and compared to their theoretical prediction according to the Standard Model. The predictions are free of form-factor uncertainties in the low  $q^2$  region. The observables  $P'_6$  and  $P'_8$  show a good agreement with their predictions. The observable  $P'_4$  shows some discrepancy in the high  $q^2$  region, however, this might just be a statistical fluctuation. The observable  $P'_5$  shows a discrepancy of the order of four standard deviations in the third bin compared with the Standard Model prediction. No experimental effect has been found yet that explains this deviation.





## 10. The End

In 2011, the LHCb detector collected more than  $1 \text{ fb}^{-1}$  of collision data at a centre-of-mass energy of  $7 \text{ TeV}$ . The tracking system of LHCb is of crucial importance for the analysis of this data. One of the tracking detectors, the Tracker Turicensis (TT), is a silicon strip detector with four layers and is placed upstream of the magnet. The efficiency to add information left by a charged particle in a layer of TT to tracks stemming from a particle traversing the whole LHCb detector was examined and improved. This «hit-adding» efficiency is about 97% for a broad momentum range, given that the particle left enough information in at least three layers of TT. Understanding the efficiency of the track reconstruction plays a key role in many physics analyses. To measure this track reconstruction efficiency, a novel method was developed. It uses the muon system and the TT to form so-called «MuonTT» tracks which allow to test the efficiency to reconstruct a long track. The track reconstruction efficiency for high transverse momentum tracks is in the range 98% - 99%, the efficiency for low transverse momentum tracks around 96% - 97%.

The decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  is an excellent laboratory for searches for physics beyond the Standard Model. The zero-crossing point of the forward-backward asymmetry of the dimuon system,  $A_{\text{FB}}$ , can be precisely predicted due to cancellation of form-factor uncertainties. It was measured for the first time in an experiment, using an «unbinned counting» technique which only requires little a priori assumptions about the shape of  $A_{\text{FB}}$ . The experimental value is

$$q_0^2 = (4.9 \pm 0.9) \text{ GeV}^2/c^4,$$

which is in good agreement with theoretical predictions. In addition to the zero-crossing point, four new observables were measured in  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  which have a low theoretical uncertainty in the region of low invariant dimuon mass. They are called  $P'_4, P'_5, P'_6$  and  $P'_8$ .  $P'_6$  and  $P'_8$  agree well with SM predictions, while  $P'_4$  shows a small tension for high invariant dimuon mass, which, however, is statistically not significant. The observable  $P'_5$  shows a discrepancy on the order of four standard deviations in the region of low invariant dimuon mass. This deviation could not be explained experimentally so far.



## A. Linear track fit

Muon standalone tracks are fitted with a linear function in the  $x$ - $z$ - and the  $y$ - $z$ -projection. The derivation for the fit in the  $x$ - $z$ -projection is given here. The linear function is of the form:

$$x(z) = a \cdot z + b \quad (\text{A.1})$$

The parameters  $a$  and  $b$  are determined using a  $\chi^2$ -minimisation. The  $\chi^2$  is defined as:

$$\chi^2 = \sum_i \frac{(x_{\text{hit},i} - x_{\text{fit},i})^2}{\sigma_{\text{hit},i}^2} \quad (\text{A.2})$$

$$\rightarrow \chi^2 = \sum_i \frac{(x_{\text{hit},i} - (a \cdot z_{\text{hit},i} + b))^2}{\sigma_{\text{hit},i}^2}, \quad (\text{A.3})$$

where  $x_{\text{hit},i}$  is the  $x$ -position of the  $i$ th hit,  $x_{\text{fit},i}$  the  $x$ -position of the linear fit at the  $z$ -position of the  $i$ th hit and  $z_{\text{hit},i}$  the  $z$ -position of the  $i$ th hit.  $\sigma_{\text{hit},i}$  is the uncertainty on the  $x$ -position of the  $i$ th hit.

The best estimators for  $a$  and  $b$ , assuming Gaussian errors on the hit-positions, can be found by minimising the  $\chi^2$  with respect to both parameters. This yields the following two equations:

$$\frac{\partial \chi^2}{\partial a} = -2 \cdot \sum_i \frac{x_{\text{hit},i} \cdot z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} + 2a \cdot \sum_i \frac{z_{\text{hit},i}^2}{\sigma_{\text{hit},i}^2} \quad (\text{A.4})$$

$$+ 2b \cdot \sum_i \frac{z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} \stackrel{!}{=} 0 \quad (\text{A.5})$$

$$\frac{\partial \chi^2}{\partial b} = -2 \cdot \sum_i \frac{x_{\text{hit},i}}{\sigma_{\text{hit},i}^2} + 2a \cdot \sum_i \frac{z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} + 2b \cdot \sum_i \frac{1}{\sigma_{\text{hit},i}^2} \stackrel{!}{=} 0 \quad (\text{A.6})$$

$$\rightarrow \sum_i \frac{x_{\text{hit},i} \cdot z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} = a \cdot \sum_i \frac{z_{\text{hit},i}^2}{\sigma_{\text{hit},i}^2} + b \cdot \sum_i \frac{z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} \quad (\text{A.7})$$

$$\rightarrow \sum_i \frac{x_{\text{hit},i}}{\sigma_{\text{hit},i}^2} = a \cdot \sum_i \frac{z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} + b \cdot \sum_i \frac{1}{\sigma_{\text{hit},i}^2} \quad (\text{A.8})$$

The last two equations can be written in a matrix form:

$$\begin{pmatrix} S_1 & S_z \\ S_z & S_{zz} \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} S_x \\ S_{xz} \end{pmatrix} \quad (\text{A.9})$$

with

$$\begin{aligned} S_1 &= \sum_i \frac{1}{\sigma_{\text{hit},i}^2} & S_z &= \sum_i \frac{z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} & S_{zz} &= \sum_i \frac{z_{\text{hit},i}^2}{\sigma_{\text{hit},i}^2} \\ S_x &= \sum_i \frac{x_{\text{hit},i}}{\sigma_{\text{hit},i}^2} & S_{xz} &= \sum_i \frac{x_{\text{hit},i} \cdot z_{\text{hit},i}}{\sigma_{\text{hit},i}^2} & S_{xx} &= \sum_i \frac{x_{\text{hit},i}^2}{\sigma_{\text{hit},i}^2} \end{aligned} \quad (\text{A.10})$$

$S_{xx}$  will be used later. Eq. A.9 can then be solved with matrix inversion:

$$\begin{pmatrix} S_1 & S_z \\ S_z & S_{zz} \end{pmatrix}^{-1} = \frac{1}{D} \cdot \begin{pmatrix} S_{zz} & -S_z \\ -S_z & S_1 \end{pmatrix}, \quad (\text{A.11})$$

with  $D$  the determinant of the matrix:  $D = S_1 S_{zz} - S_z^2$ . The inverted matrix is the same as the covariance matrix. The solutions for the best estimators can be derived as:

$$\hat{a} = \frac{1}{D} (S_{zz} S_x - S_z S_{xz}) \quad (\text{A.12})$$

$$\hat{b} = \frac{1}{D} (-S_z S_x + S_1 S_{xz}). \quad (\text{A.13})$$

with the uncertainties:

$$\sigma_a = \sqrt{\frac{S_1}{D}} \quad (\text{A.14})$$

$$\sigma_b = \sqrt{\frac{S_{zz}}{D}} \quad (\text{A.15})$$

The goodness-of-fit can be calculated as:

$$\chi^2/\text{ndof} = \frac{1}{\text{ndof}} \cdot \left( S_{xx} + \hat{a}^2 S_{zz} + \hat{b}^2 S_1 - 2\hat{a} S_{xz} - 2\hat{b} S_x + \hat{a}\hat{b} S_z \right) \quad (\text{A.16})$$

## B. Calculation of uncertainty with sideband subtraction

The following problem often occurs when determining the efficiency with a tag-and-probe method. One has two invariant mass distributions with a peak, for example  $J/\psi \rightarrow \mu^+ \mu^-$ , where for one distribution the probe has to fulfill a certain condition while for the other distribution no such requirement is made. To determine the efficiency for this requirement, the number of signal events in both distributions has to be counted. If the distribution is not background free but the background has a linear shape, a sideband subtraction can be performed. The number of signal events equals the number of events in a signal window minus the number of events in two sidebands, where the width of the two sidebands together is the same as the width of the signal window. Furthermore, the distance between the left sideband and the signal window has to be the same as the distance between the right sideband and the signal window. Such a situation is depicted in Fig. 6.6.

The efficiency is then:

$$\varepsilon = \frac{\# \text{ number of candidates where probe fulfills condition}}{\# \text{ number of candidates}} \quad (\text{B.1})$$

The uncertainty on the efficiency can then be calculated in the following way. Let  $k_{\text{sig}}$  and  $k_{\text{SB}}$  be the number of candidates that fulfill the requirement in the signal window and in the sidebands respectively, and let  $f_{\text{sig}}$  and  $f_{\text{SB}}$  be the number of candidates that fail the requirement in the signal window and in the sidebands. The total number of candidates in the signal window and the sidebands,  $N_{\text{sig}}$  and  $N_{\text{SB}}$ , is then  $N_{\text{sig}} = k_{\text{sig}} + f_{\text{sig}}$  and  $N_{\text{SB}} = k_{\text{SB}} + f_{\text{SB}}$ . The efficiency in this notation is:

$$\varepsilon = \frac{k_{\text{sig}} - k_{\text{SB}}}{k_{\text{sig}} + f_{\text{sig}} - (k_{\text{SB}} + f_{\text{SB}})} \quad (\text{B.2})$$

The uncertainty can then be derived via standard error propagation:

$$\sigma_{\varepsilon}^2 = \left( \frac{\partial \varepsilon}{\partial k_{\text{sig}}} \right)^2 \sigma_{k_{\text{sig}}}^2 + \left( \frac{\partial \varepsilon}{\partial k_{\text{SB}}} \right)^2 \sigma_{k_{\text{SB}}}^2 + \left( \frac{\partial \varepsilon}{\partial f_{\text{sig}}} \right)^2 \sigma_{f_{\text{sig}}}^2 + \left( \frac{\partial \varepsilon}{\partial f_{\text{SB}}} \right)^2 \sigma_{f_{\text{SB}}}^2 \quad (\text{B.3})$$

Using the substitution  $k_{\text{sig, SB}} = \varepsilon N_{\text{sig, SB}}$  and  $\sigma_{k_{\text{sig, SB}}} = \sqrt{k_{\text{sig, SB}}}$ ,  $\sigma_{f_{\text{sig, SB}}} = \sqrt{f_{\text{sig, SB}}}$  leads to:

$$\sigma_\varepsilon^2 = \frac{(1 - \varepsilon)^2(k_{\text{sig}} + k_{\text{SB}}) + \varepsilon^2(f_{\text{sig}} + f_{\text{SB}})}{N_{\text{sig}} - N_{\text{SB}}} \quad (\text{B.4})$$

$$= \frac{(1 - 2\varepsilon)(k_{\text{sig}} + k_{\text{SB}}) + \varepsilon^2(N_{\text{sig}} + N_{\text{SB}})}{N_{\text{sig}} - N_{\text{SB}}}. \quad (\text{B.5})$$

In the last line it was used that  $f_{\text{sig,SB}} = N_{\text{sig,SB}} - k_{\text{sig,SB}}$ . Equation B.5 can now be used to calculate the uncertainty when determined with a sideband subtraction.

## C. Extraction of a reference value of the zero-crossing point of $A_{\text{FB}}$ in simulation

As the zero-crossing point is not a parameter of the  $A_{\text{FB}}$  distribution of the Standard Model, it had to be determined from the  $A_{\text{FB}}$  distribution of the simulated sample itself. The full simulated sample was into 50 bins in the range from 0  $\text{GeV}^2/c^4$  and 6  $\text{GeV}^2/c^4$ . For each bin  $A_{\text{FB}}$  was calculated and the distribution was fitted with a linear function from 3  $\text{GeV}^2/c^4$  to 5  $\text{GeV}^2/c^4$ . This is shown in Fig. C.1. The extracted zero-crossing point is 4.07  $\text{GeV}^2/c^4$  with an uncertainty of less than 1%.

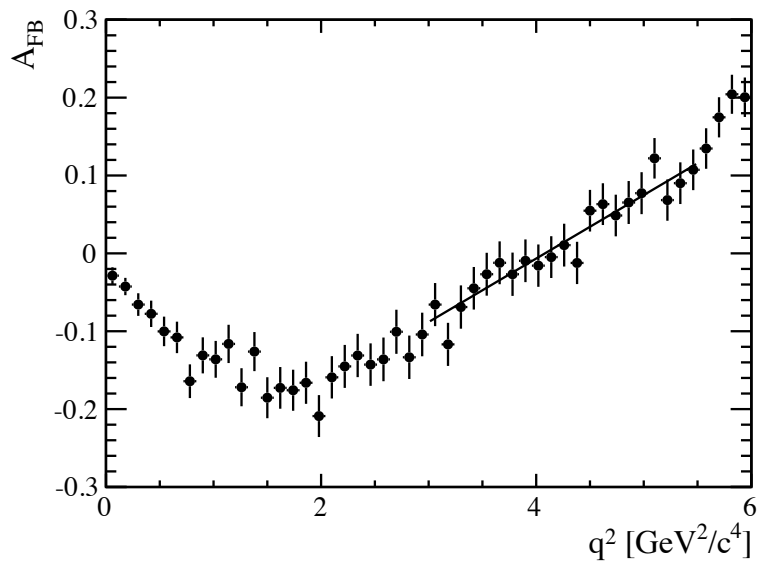


Figure C.1.:  $A_{\text{FB}}$  distribution in simulated sample. A linear function was fitted in the range  $3 \text{ GeV}^2/c^4$  to  $5 \text{ GeV}^2/c^4$  to extract the zero-crossing point.



## D. Derivation of formulae for the $S_i$ and $P'_i$ observables

The differential decay rate for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ , neglecting the lepton masses and an S-wave contribution, is:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4\Gamma}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{32\pi} \left( I_1^s \sin^2 \theta_K + I_1^c \cos^2 \theta_K + (I_2^s \sin^2 \theta_K + I_2^c \cos^2 \theta_K) \cos 2\theta_\ell \right. \\ & + I_3 \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \\ & + I_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \\ & + (I_6^s \sin^2 \theta_K + I_6^c \cos^2 \theta_K) \cos \theta_\ell + I_7 \sin 2\theta_K \sin \theta_\ell \sin \phi \\ & \left. + I_8 \sin 2\theta_K \sin 2\theta_\ell \sin \phi + I_9 \sin^2 \theta_K \sin^2 \theta_\ell \sin 2\phi \right) \end{aligned}$$

For the subsequent calculations, some considerations are helpful. The first are the formulae for the sin and cos of the sum of two angles:

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \sin \beta \cos \alpha \quad (\text{D.1})$$

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \quad (\text{D.2})$$

The behaviour of  $\sin \alpha$ ,  $\cos \alpha$ ,  $\sin 2\alpha$  and  $\cos 2\alpha$  terms under certain transformations is summarised in Table D.1.

When folding a distribution using one of the relations in Table D.1, the terms transforming with a minus sign will therefore cancel. In total there are three cases:

- a) The term does not contain an angle which is folded: The term is multiplied with a factor of two to keep the normalisation correct.
- b) The term does contain an angle which is folded and the angle is symmetric under the transformation: The term is multiplied with a factor of two as the distributions add up.
- c) The term does contain an angle which is folded and the angle is antisymmetric under the transformation: The term cancels.

Table D.1.: Properties of angular expressions under the transformations used to simplify the PDFs for the  $S_i$  and  $P'_i$  observables.

Expression	$\alpha \rightarrow -\alpha$	$\alpha \rightarrow \pi - \alpha$	$\alpha \rightarrow -\pi - \alpha$
$\cos \alpha$	$\cos \alpha$	$-\cos \alpha$	$-\cos \alpha$
$\cos 2\alpha$	$\cos 2\alpha$	$\cos 2\alpha$	$\cos 2\alpha$
$\sin \alpha$	$-\sin \alpha$	$\sin \alpha$	$\sin \alpha$
$\sin 2\alpha$	$-\sin 2\alpha$	$-\sin 2\alpha$	$-\sin 2\alpha$

As in case a) and b) the terms are multiplied with a factor of two, the factor can be neglected, as the overall normalisation is not of importance for the fitting procedure, only the relative one between the different terms.

Furthermore the following relations are used:

$$I_1^s = \frac{3}{4}(1 - F_L) \quad (\text{D.3})$$

$$I_1^c = F_L \quad (\text{D.4})$$

$$I_2^s = \frac{1}{4}(1 - F_L) \quad (\text{D.5})$$

$$I_2^c = -F_L, \quad (\text{D.6})$$

and implicitly also

$$S_3 = \frac{1}{2}(1 - F_L)A_T^{(2)} \quad (\text{D.7})$$

is used.

## D.1. Transformations for $P'_4/S_4$

First the  $\phi$  angle is folded with  $\phi \rightarrow \phi$  for  $\phi < 0$ . This leads to the differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4\Gamma}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} = & \frac{9}{16\pi} \left( I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + (I_2^s \sin^2\theta_K + I_2^c \cos^2\theta_K) \cos 2\theta_\ell \right. \\ & + I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \\ & + I_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \\ & \left. + (I_6^s \sin^2\theta_K + I_6^c \cos^2\theta_K) \cos \theta_\ell \right) \end{aligned}$$

Then the angle  $\phi$  is folded again, together with  $\theta_\ell$ , following the rule:  $\phi \rightarrow \pi - \phi$  and  $\theta_\ell \rightarrow \pi - \theta_\ell$  if  $\theta_\ell > \frac{\pi}{2}$ . Also adding the  $\mathcal{CP}$  conjugate decay, this then leads to:

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{3}{4}(1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ & + \frac{1}{4}(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2}(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & \left. + \sqrt{F_L(1 - F_L)} P'_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \right) \end{aligned}$$

## D.2. Transformations for $P'_5/S_5$

First the  $\phi$  angle is folded with  $\phi \rightarrow \phi$  for  $\phi < 0$ . This leads to the differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4\Gamma}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{16\pi} \left( I_1^s \sin^2 \theta_K + I_1^c \cos^2 \theta_K + (I_2^s \sin^2 \theta_K + I_2^c \cos^2 \theta_K) \cos 2\theta_\ell \right. \\ & + I_3 \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \\ & + I_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \\ & \left. + (I_6^s \sin^2 \theta_K + I_6^c \cos^2 \theta_K) \cos \theta_\ell \right) \end{aligned}$$

Then the angle  $\theta_\ell$  is folded, following the rule:  $\theta_\ell \rightarrow \pi - \theta_\ell$  if  $\theta_\ell > \frac{\pi}{2}$ . Also adding the  $\mathcal{CP}$  conjugate decay, this then leads to:

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{3}{4}(1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ & + \frac{1}{4}(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\ & + \frac{1}{2}(1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ & \left. + \sqrt{F_L(1 - F_L)} P'_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \right) \end{aligned}$$

## D.3. Transformations for $P'_6/S_7$

First the  $\phi$  angle is folded with  $\phi \rightarrow \pi - \phi$  for  $\phi > \frac{\pi}{2}$  and  $\phi \rightarrow -\pi - \phi$  for  $\phi < -\frac{\pi}{2}$ . This leads to the differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4\Gamma}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} &= \frac{9}{16\pi} \left( I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + (I_2^s \sin^2\theta_K + I_2^c \cos^2\theta_K) \cos 2\theta_\ell \right. \\ &\quad + I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi \\ &\quad + I_5 \sin 2\theta_K \sin\theta_\ell \cos\phi \\ &\quad + (I_6^s \sin^2\theta_K + I_6^c \cos^2\theta_K) \cos\theta_\ell \\ &\quad \left. + I_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + I_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi \right) \end{aligned}$$

Then the angle  $\theta_\ell$  is folded, following the rule:  $\theta_\ell \rightarrow \pi - \theta_\ell$  if  $\theta_\ell > \frac{\pi}{2}$ . Adding also the  $\mathcal{CP}$  conjugate decay, this then leads to:

$$\begin{aligned} \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} &= \frac{9}{8\pi} \left( \frac{3}{4}(1 - F_L) \sin^2\theta_K + F_L \cos^2\theta_K \right. \\ &\quad + \frac{1}{4}(1 - F_L) \sin^2\theta_K \cos 2\theta_\ell - F_L \cos^2\theta_K \cos 2\theta_\ell \\ &\quad + \frac{1}{2}(1 - F_L) A_T^{(2)} \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi \\ &\quad \left. + \sqrt{F_L(1 - F_L)} P'_6 \sin 2\theta_K \sin\theta_\ell \sin\phi \right) \end{aligned}$$

## D.4. Transformations for $P'_8/S_8$

First the  $\phi$  angle is folded with  $\phi \rightarrow \pi - \phi$  for  $\phi > \frac{\pi}{2}$  and  $\phi \rightarrow -\pi - \phi$  for  $\phi < -\frac{\pi}{2}$ . This leads to the differential decay rate:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d^4\Gamma}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} &= \frac{9}{16\pi} \left( I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + (I_2^s \sin^2\theta_K + I_2^c \cos^2\theta_K) \cos 2\theta_\ell \right. \\ &\quad + I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi \\ &\quad + I_5 \sin 2\theta_K \sin\theta_\ell \cos\phi \\ &\quad + (I_6^s \sin^2\theta_K + I_6^c \cos^2\theta_K) \cos\theta_\ell \\ &\quad \left. + I_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + I_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi \right) \end{aligned}$$

Then the angles  $\theta_\ell$  and  $\theta_K$  are folded, following the rule:  $\theta_\ell \rightarrow \pi - \theta_\ell$  and  $\theta_K \rightarrow \pi - \theta_K$  if  $\theta_\ell > \frac{\pi}{2}$ . Adding also the  $\mathcal{CP}$  conjugate decay this then leads to:

$$\begin{aligned}
\frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{d \cos \theta_\ell d \cos \theta_K d\phi dq^2} = & \frac{9}{8\pi} \left( \frac{3}{4} (1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\
& + \frac{1}{4} (1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell - F_L \cos^2 \theta_K \cos 2\theta_\ell \\
& + \frac{1}{2} (1 - F_L) A_T^{(2)} \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\
& \left. + \sqrt{F_L(1 - F_L)} P'_8 \sin 2\theta_K \sin 2\theta_\ell \cos \phi \right)
\end{aligned}$$



## E. S-wave contribution to

$$B^0 \rightarrow K^{*0} \mu^+ \mu^-$$

The introduction of an S-wave modifies the angular observables for  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ . Using the convention of Refs. [139] and [140] the S-wave angular distribution is the following:

$$W_s = \frac{1}{4\pi} \left( \tilde{I}_{1a} + \tilde{I}_{1b} \cos \theta_K + (\tilde{I}_{2a} + \tilde{I}_{2b} \cos \theta_K) \cos 2\theta_l + \right. \\ \left. \tilde{I}_4 \sin \theta_K \sin 2\theta_l \cos \phi + \tilde{I}_5 \sin \theta_K \sin \theta_l \cos \phi + \right. \\ \left. \tilde{I}_7 \sin \theta_K \sin \theta_l \sin \phi + \tilde{I}_8 \sin \theta_K \sin 2\theta_l \sin \phi \right) \quad (\text{E.1})$$

When neglecting the lepton masses, one has:

$$\tilde{I}_{1a} = -\tilde{I}_{2a} \quad (\text{E.2})$$

$$\tilde{I}_{1b} = -\tilde{I}_{2b} \quad (\text{E.3})$$

and therefore  $W_s$  becomes.

$$W_s = \frac{1}{4\pi} \left( 2\tilde{I}_{1a} \sin^2 \theta_l + 2\tilde{I}_{1b} \sin^2 \theta_l \cos \theta_K + \right. \\ \left. \tilde{I}_4 \sin \theta_K \sin 2\theta_l \cos \phi + \tilde{I}_5 \sin \theta_K \sin \theta_l \cos \phi + \right. \\ \left. \tilde{I}_7 \sin \theta_K \sin \theta_l \sin \phi + \tilde{I}_8 \sin \theta_K \sin 2\theta_l \sin \phi \right) \quad (\text{E.4})$$

After some calculations and when writing the full angular distribution in the form

$$\frac{d^4\Gamma}{d\cos\theta_\ell d\cos\theta_K d\phi dq^2} \propto (1 - F_S) PDF_{K^{*0}} + PDF_S \quad (\text{E.5})$$

the S-wave contribution can be written as:

$$PDF_S = \frac{2}{3} F_S \sin^2 \theta_\ell + \frac{4}{3} A_S \sin^2 \theta_\ell \cos \theta_K + \\ A_S^{(4)} \sin \theta_K \sin 2\theta_\ell \cos \phi + \\ A_S^{(5)} \sin \theta_K \sin \theta_\ell \cos \phi + \\ A_S^{(7)} \sin \theta_K \sin \theta_\ell \sin \phi + \\ A_S^{(8)} \sin \theta_K \sin 2\theta_\ell \sin \phi. \quad (\text{E.6})$$

with:

$$F_S = \frac{A_{0,S} A_{0,S}^*}{\Gamma_{tot}} = \frac{\Gamma_S}{\Gamma_{K^*} + \Gamma_S} \quad (\text{E.7})$$

$$A_S = \sqrt{3} \frac{\text{Re} \left( A_0^{*L,R} A_0^{0(L,R)} \right)}{\Gamma_{tot}} \quad (\text{E.8})$$

$$A_s^{(4)} = \sqrt{\frac{2}{3}} \frac{\text{Re} \left( A_{\parallel}^{*L} A_0^{0(L)} + A_{\parallel}^{*R} A_0^{0(R)} \right)}{\Gamma_{tot}} \quad (\text{E.9})$$

$$A_s^{(5)} = \frac{2\sqrt{2}}{\sqrt{3}} \frac{\text{Re} \left( A_{\perp}^{*L} A_0^{0(L)} - A_{\perp}^{*R} A_0^{0(R)} \right)}{\Gamma_{tot}} \quad (\text{E.10})$$

$$A_s^{(7)} = \frac{2\sqrt{2}}{\sqrt{3}} \frac{\text{Re} \left( A_{\parallel}^{*L} A_0^{0(L)} - A_{\parallel}^{*R} A_0^{0(R)} \right)}{\Gamma_{tot}} \quad (\text{E.11})$$

$$A_s^{(8)} = \sqrt{\frac{2}{3}} \frac{\text{Re} \left( A_{\perp}^{*L} A_0^{0(L)} + A_{\perp}^{*R} A_0^{0(R)} \right)}{\Gamma_{tot}} \quad (\text{E.12})$$



## F. Pulls for $A_T^{(2)}$ and $F_L$ for $P'_4$

The Figs. F.1 and F.2 show the mean and the RMS of the pulls for  $A_T^{(2)}$  and  $F_L$  obtained with a toy study.

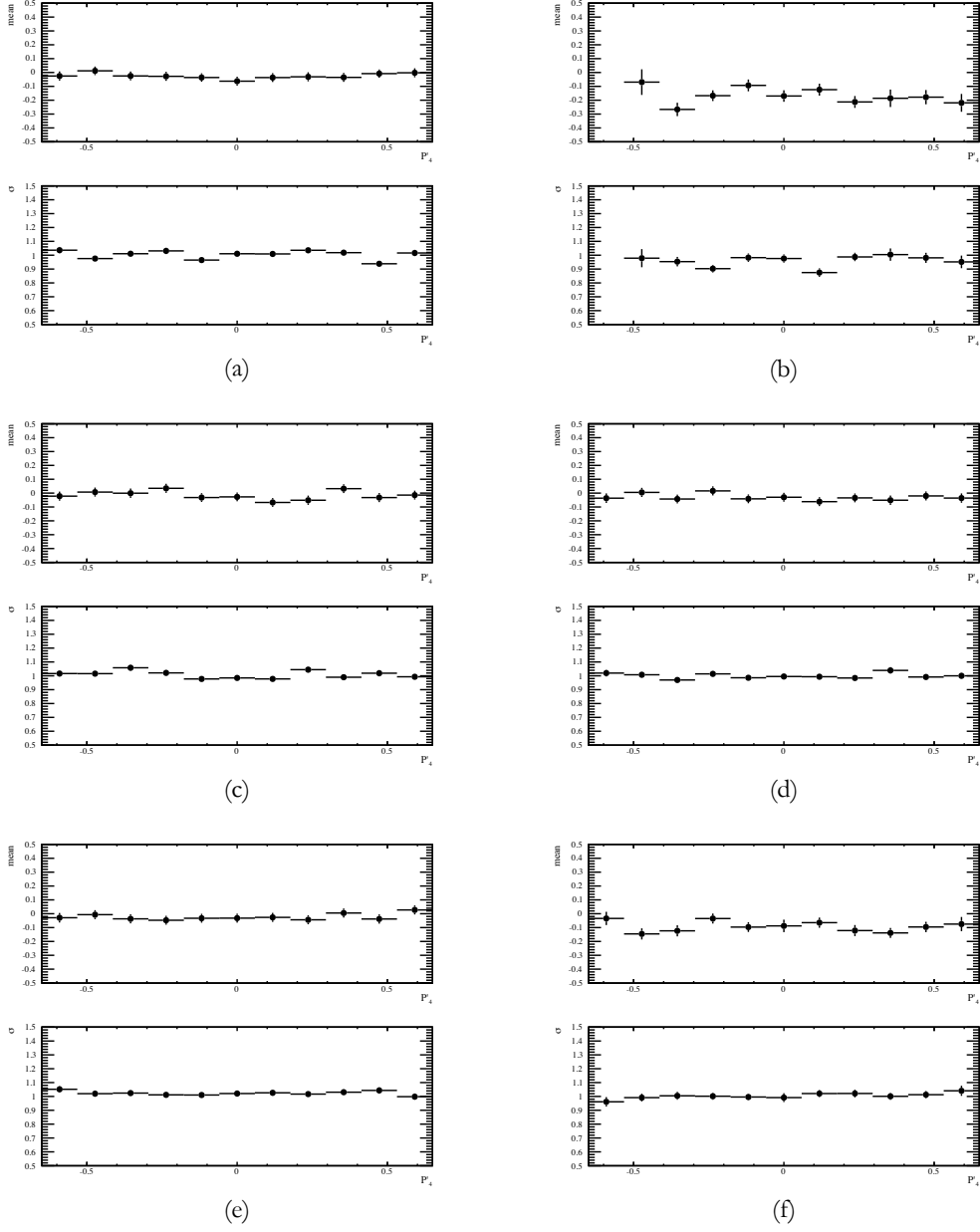
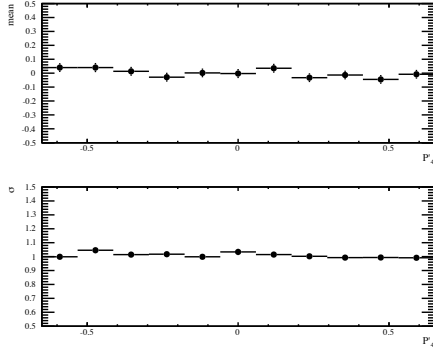
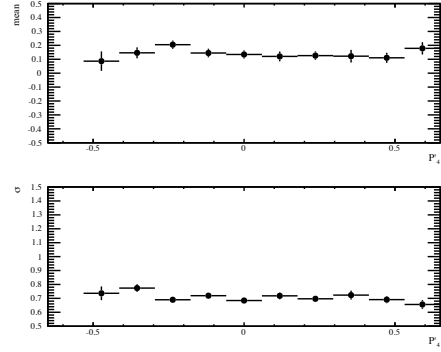


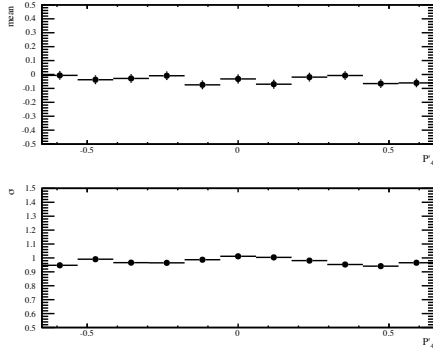
Figure F.1.: Mean and RMS values for  $F_L$  for the pulls from toy samples of  $P_4'$  in all six  $q^2$  bins for 11 possible  $P_4'$  values in each bin. Except for the regions close to the boundary in bin 2 (b), the deviation is always smaller than 10%.



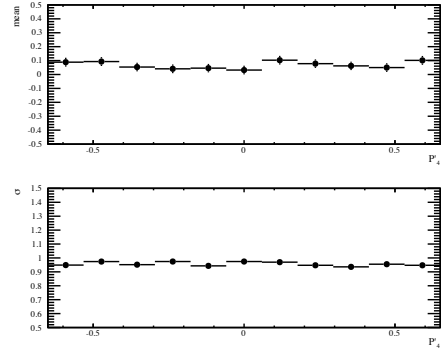
(a)



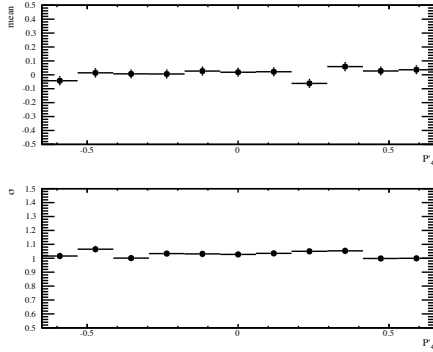
(b)



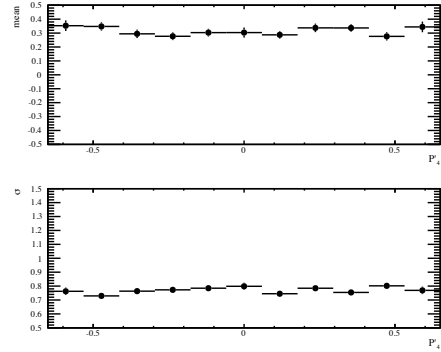
(c)



(d)



(e)



(f)

Figure F.2.: Mean and RMS values for  $A_T^{(2)}$  for the pulls from toy samples of  $P_4'$  in all six  $q^2$  bins for 11 possible  $P_4'$  values in each bin. Except for the regions close to the boundary in bin 2 (b), the deviation is always smaller than 10%.



## G. Systematic uncertainties for $P'_4$ , $P'_5$ , $P'_6$ , $P'_8$

The Tables G.1, G.2, G.3 and G.4 list the systematic uncertainties for the  $P'_4$ ,  $P'_5$ ,  $P'_6$  and  $P'_8$  observables. The definition of the categories is given in Sect. 8.6.

Table G.1.: Systematic uncertainties for  $P'_4$ . The definition of the categories is given in Sect. 8.6.

Type	Bin 1	Deviation	Bin 2	Deviation	Bin 3	Deviation	Bin 4	Deviation	Bin 5	Deviation	Bin 6	Deviation
1	0.438	0.015	0.288	0.004	-0.192	0.004	-0.786	0.012	-0.787	0.006	-0.621	0.016
2	0.451	0.002	0.295	0.003	-0.186	0.002	-0.813	0.015	-0.795	0.002	-0.608	0.003
3	0.449	0.003	0.294	0.003	-0.187	0.001	-0.795	0.004	-0.79	0.003	-0.613	0.008
4	0.453	0.0	0.29	0.002	-0.187	0.002	-0.801	0.003	-0.79	0.002	-0.603	0.002
5	0.453	0.0	0.29	0.002	-0.187	0.002	-0.801	0.003	-0.79	0.002	-0.603	0.002
6	0.453	0.0	0.29	0.002	-0.187	0.002	-0.801	0.003	-0.79	0.002	-0.603	0.002
7	0.452	0.001	0.287	0.004	-0.184	0.004	-0.788	0.011	-0.787	0.006	-0.602	0.002
8	0.451	0.002	0.291	0.001	-0.187	0.002	-0.79	0.008	-0.791	0.001	-0.603	0.002
9	0.455	0.002	0.292	0.0	-0.189	0.001	-0.793	0.005	-0.793	0.0	-0.606	0.001
10	0.455	0.002	0.292	0.0	-0.189	0.001	-0.793	0.005	-0.793	0.0	-0.606	0.001

Table G.2.: Systematic uncertainties for  $P'_5$ . The definition of the categories is given in Sect. 8.6.

Type	Bin 1	Deviation	Bin 2	Deviation	Bin 3	Deviation	Bin 4	Deviation	Bin 5	Deviation	Bin 6	Deviation
1	0.438	0.015	0.288	0.004	-0.192	0.004	-0.786	0.012	-0.787	0.006	-0.621	0.016
2	0.451	0.002	0.295	0.003	-0.186	0.002	-0.813	0.015	-0.795	0.002	-0.608	0.003
	0.449	0.003	0.294	0.003	-0.187	0.001	-0.795	0.004	-0.79	0.003	-0.613	0.008
3	0.453	0.0	0.294	0.002	-0.189	0.0	-0.805	0.007	-0.793	0.0	-0.616	0.01
	0.452	0.001	0.29	0.001	-0.188	0.0	-0.789	0.009	-0.793	0.0	-0.615	0.011
4	0.453	0.0	0.29	0.002	-0.187	0.002	-0.801	0.003	-0.79	0.002	-0.603	0.002
	0.453	0.0	0.29	0.002	-0.187	0.002	-0.801	0.003	-0.79	0.002	-0.603	0.002
5	0.45	0.003	0.291	0.0	-0.188	0.0	-0.797	0.001	-0.792	0.0	-0.609	0.004
	0.455	0.002	0.292	0.0	-0.189	0.001	-0.793	0.005	-0.793	0.0	-0.606	0.00
	0.456	0.003	0.298	0.006	-0.187	0.002	-0.787	0.012	-0.783	0.01	-0.618	0.013
	0.45	0.003	0.286	0.005	-0.19	0.001	-0.815	0.016	-0.803	0.01	-0.619	0.014
	0.456	0.003	0.296	0.005	-0.184	0.004	-0.797	0.002	-0.784	0.008	-0.603	0.002
6	0.459	0.006	0.303	0.012	-0.182	0.006	-0.778	0.021	-0.774	0.019	-0.613	0.008
	0.447	0.005	0.281	0.01	-0.194	0.006	-0.823	0.024	-0.81	0.017	-0.613	0.009
	0.45	0.003	0.287	0.005	-0.193	0.004	-0.808	0.009	-0.801	0.008	-0.617	0.013
	0.446	0.007	0.256	0.035	-0.205	0.017	-0.798	0.0	-0.79	0.003	-0.597	0.007
	0.461	0.008	0.329	0.038	-0.171	0.017	-0.799	0.001	-0.796	0.003	-0.612	0.008
7	0.453	0.00	0.279	0.012	-0.185	0.003	-0.806	0.008	-0.787	0.006	-0.601	0.004
	0.453	0.00	0.305	0.014	-0.192	0.004	-0.794	0.005	-0.798	0.005	-0.609	0.004
8	0.452	0.001	0.287	0.004	-0.184	0.004	-0.788	0.011	-0.787	0.006	-0.602	0.002
9	0.451	0.002	0.291	0.001	-0.187	0.002	-0.79	0.008	-0.791	0.001	-0.603	0.002
10	0.455	0.002	0.292	0.0	-0.189	0.001	-0.793	0.005	-0.793	0.0	-0.606	0.001

Table G.3.: Systematic uncertainties for  $P'_6$ . The definition of the categories is given in Sect. 8.6.

Type	Bin 1	Deviation	Bin 2	Deviation	Bin 3	Deviation	Bin 4	Deviation	Bin 5	Deviation	Bin 6	Deviation
1	-0.239	0.002	0.156	0.009	-0.045	0.001	-0.319	0.005	-0.179	0.001	0.304	0.002
2	-0.234	0.003	0.155	0.008	-0.045	0.001	-0.308	0.007	-0.179	0.001	0.314	0.007
3	-0.237	0.0	0.154	0.007	-0.043	0.002	-0.314	0.0	-0.183	0.003	0.307	0.001
4	-0.235	0.002	0.146	0.001	-0.045	0.001	-0.313	0.001	-0.178	0.002	0.307	0.0
5	-0.238	0.001	0.146	0.001	-0.043	0.001	-0.315	0.001	-0.179	0.001	0.308	0.001
6	-0.238	0.002	0.156	0.009	-0.043	0.002	-0.316	0.002	-0.182	0.001	0.306	0.0
7	-0.238	0.002	0.156	0.009	-0.043	0.002	-0.316	0.002	-0.182	0.001	0.306	0.0
8	-0.23	0.006	0.139	0.008	-0.044	0.001	-0.313	0.001	-0.178	0.003	0.308	0.00
9	-0.241	0.004	0.152	0.004	-0.045	0.0	-0.314	0.0	-0.179	0.001	0.307	0.001
10	-0.241	0.004	0.154	0.007	-0.042	0.002	-0.316	0.002	-0.181	0.001	0.306	0.0
11	-0.233	0.003	0.143	0.004	-0.047	0.002	-0.313	0.001	-0.18	0.0	0.307	0.0
12	-0.24	0.004	0.15	0.003	-0.04	0.005	-0.313	0.001	-0.172	0.008	0.307	0.001
13	-0.244	0.008	0.158	0.011	-0.038	0.007	-0.315	0.001	-0.173	0.008	0.307	0.0
14	-0.23	0.007	0.141	0.006	-0.051	0.006	-0.314	0.0	-0.187	0.007	0.306	0.001
15	-0.233	0.003	0.144	0.003	-0.049	0.004	-0.315	0.001	-0.187	0.007	0.305	0.001
16	-0.259	0.022	0.142	0.005	-0.035	0.01	-0.321	0.007	-0.179	0.001	0.322	0.016
17	-0.214	0.023	0.151	0.004	-0.056	0.011	-0.305	0.009	-0.185	0.005	0.293	0.013
18	-0.239	0.002	0.146	0.001	-0.041	0.004	-0.312	0.002	-0.175	0.005	0.306	0.00
19	-0.235	0.002	0.148	0.001	-0.049	0.004	-0.317	0.003	-0.188	0.008	0.307	0.001
20	-0.24	0.003	0.163	0.015	-0.041	0.003	-0.319	0.005	-0.184	0.004	0.306	0.0
21	-0.238	0.001	0.152	0.004	-0.043	0.001	-0.315	0.001	-0.181	0.001	0.306	0.0
22	-0.241	0.004	0.152	0.004	-0.045	0.0	-0.314	0.0	-0.179	0.001	0.307	0.001



Table G.4.: Systematic uncertainties for  $P'_g$ . The definition of the categories is given in Sect. 8.6.

Type	Bin 1	Deviation	Bin 2	Deviation	Bin 3	Deviation	Bin 4	Deviation	Bin 5	Deviation	Bin 6	Deviation
1	-0.067	0.006	-0.161	0.015	0.291	0.001	-0.048	0.01	-0.204	0.004	0.062	0.001
	-0.06	0.002	-0.155	0.009	0.291	0.002	-0.06	0.001	-0.204	0.004	0.058	0.002
	-0.068	0.006	-0.156	0.01	0.294	0.001	-0.057	0.002	-0.198	0.002	0.063	0.002
2	-0.059	0.003	-0.147	0.001	0.292	0.001	-0.059	0.0	-0.201	0.0	0.062	0.001
	-0.063	0.001	-0.147	0.001	0.294	0.001	-0.059	0.0	-0.2	0.0	0.061	0.0
	-0.063	0.001	-0.141	0.005	0.293	0.001	-0.059	0.0	-0.198	0.002	0.063	0.002
3	-0.063	0.001	-0.141	0.005	0.293	0.001	-0.059	0.0	-0.198	0.002	0.063	0.002
	-0.058	0.004	-0.135	0.011	0.291	0.001	-0.06	0.001	-0.2	0.0	0.06	0.001
	-0.062	0.001	-0.155	0.009	0.294	0.001	-0.058	0.0	-0.201	0.001	0.063	0.002
4	-0.064	0.002	-0.15	0.004	0.296	0.003	-0.064	0.005	-0.199	0.001	0.056	0.005
	-0.059	0.002	-0.143	0.003	0.29	0.002	-0.054	0.005	-0.202	0.001	0.066	0.005
	-0.063	0.001	-0.141	0.005	0.295	0.002	-0.062	0.003	-0.198	0.003	0.054	0.007
5	-0.065	0.004	-0.144	0.002	0.297	0.005	-0.067	0.008	-0.196	0.004	0.049	0.012
	-0.058	0.004	-0.148	0.002	0.289	0.004	-0.052	0.007	-0.204	0.004	0.072	0.011
	-0.06	0.001	-0.151	0.006	0.291	0.001	-0.056	0.003	-0.203	0.002	0.068	0.007
6	-0.074	0.012	-0.184	0.038	0.31	0.017	-0.069	0.01	-0.198	0.003	0.084	0.023
	-0.048	0.014	-0.108	0.038	0.275	0.018	-0.048	0.011	-0.204	0.004	0.039	0.022
	-0.059	0.003	-0.133	0.013	0.295	0.002	-0.061	0.002	-0.197	0.003	0.064	0.003
7	-0.064	0.002	-0.161	0.015	0.291	0.002	-0.056	0.003	-0.205	0.005	0.059	0.002
	-0.063	0.001	-0.137	0.009	0.294	0.001	-0.057	0.001	-0.196	0.004	0.063	0.002
	-0.064	0.002	-0.139	0.007	0.293	0.001	-0.059	0.0	-0.198	0.002	0.061	0.0
8	-0.062	0.001	-0.155	0.009	0.294	0.001	-0.058	0.0	-0.201	0.001	0.063	0.002



# References

- [1] S. Weinberg, *What is an elementary particle?*, Beam Line - A periodical of particle physics (1997).
- [2] C. Quigg, *Unanswered Questions in the Electroweak Theory*, Ann. Rev. Nucl. Part. Sci. **59** (2009) 505, [arXiv:0905.3187](#).
- [3] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.
- [4] P. Skands, *Introduction to QCD*, [arXiv:1207.2389](#).
- [5] R. K. Ellis, W. J. Stirling, and B. Webber, *QCD and collider physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996) 1.
- [6] J. Horejsi, *Fundamentals of electroweak theory*. The Karolinum Press, 2002.
- [7] T. Binoth, C. Buttar, P. J. Clark, and E. Glover, eds., *LHC Physics*. CRC Press, Taylor & Francis Group, 2012.
- [8] Particle Data Group, K. Nakamura *et al.*, *Review of particle physics*, J. Phys. G **G37** (2010) 075021.
- [9] S. W. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.
- [10] C. Bebek *et al.*, *Evidence for New-Flavor Production at the  $\Upsilon(4S)$* , Phys. Rev. Lett. **46** (1981) 84.
- [11] K. Chadwick *et al.*, *Decay of b-flavored hadrons to single-muon and dimuon final states*, Phys. Rev. Lett. **46** (1981) 88.
- [12] ARGUS Collaboration, H. Albrecht *et al.*, *Observation of  $B^0$  - anti- $B^0$  Mixing*, Phys. Lett. **B192** (1987) 245.
- [13] F. Bernlochner, *Determination of the CKM matrix element  $|V_{cb}|$ , the  $B \rightarrow X s \gamma$  decay rate, and the b-quark mass*. PhD Thesis, HU-Berlin, 2011.

- [14] C. Bouchiat, J. Iliopoulos, and P. Meyer, *An anomaly-free version of Weinberg's model*, Physics Letters B **38** (1972), no. 7 519 .
- [15] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Measurements of the solar neutrino flux from Super-Kamiokande's first 300 days*, Phys. Rev. Lett. **81** (1998) 1158, [arXiv:hep-ex/9805021](#).
- [16] M. Maltoni, T. Schwetz, M. Tortola, and J. Valle, *Status of global fits to neutrino oscillations*, New J. Phys. **6** (2004) 122, [arXiv:hep-ph/0405172](#).
- [17] Particle Data Group, J. Beringer *et al.*, *Review of Particle Physics (RPP)*, Phys. Rev. **D86** (2012) 010001.
- [18] ATLAS Collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, [arXiv:1207.7214](#).
- [19] CMS Collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, [arXiv:1207.7235](#).
- [20] P. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964), no. 2 132 .
- [21] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [22] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.
- [23] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [24] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [25] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [26] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.
- [27] E. Fermi, *Versuch einer Theorie der  $\beta$ -Strahlen. I*, Zeitschrift für Physik A Hadrons and Nuclei **88** (1934) 161.
- [28] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helvetica Physica Acta **6** (1933) 110.

- [29] D. Clowe, A. Gonzalez, and M. Markevitch, *Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter*, Astrophys. J. **604** (2004) 596, [arXiv:astro-ph/0312273](#).
- [30] M. Roos, *Expansion of the Universe - Standard Big Bang Model*, [arXiv:0802.2005](#).
- [31] A. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.
- [32] J. M. Cline, *Baryogenesis*, [arXiv:hep-ph/0609145](#).
- [33] S. P. Martin, *A Supersymmetry primer*, [arXiv:hep-ph/9709356](#).
- [34] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [35] LHCb, R. Aaij *et al.*, *Measurement of  $\sigma(pp \rightarrow b\bar{b}X)$  at  $\sqrt{s}=7$  TeV in the forward region*, Phys. Lett. **B694** (2010) 209, [arXiv:1009.2731](#).
- [36] LHCb Collaboration, R. Aaij *et al.*, *Determination of  $f_s/f_d$  for 7 TeV  $pp$  collisions and a measurement of the branching fraction of the decay  $B_d \rightarrow D^- K^+$* , Phys. Rev. Lett. **107** (2011) 211801, [arXiv:1106.4435](#).
- [37] CDF Collaboration, T. Aaltonen *et al.*, *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$ -TeV*, Phys. Rev. **D77** (2008) 072003, [arXiv:0801.4375](#).
- [38] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, [arXiv:hep-ph/0005110](#).
- [39] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of  $b$ -hadron,  $c$ -hadron, and tau-lepton properties as of early 2012*, [arXiv:1207.1158](#).
- [40] L. Kowarski, *An account of the origin and beginnings of CERN*. CERN, Geneva, 1961.
- [41] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008).
- [42] L. Evans, *The Large Hadron Collider: a Marvel of Technology*. Fundamental Sciences. EPFL Press, Lausanne, 2009.
- [43] LHC Civil Engineering Webpage, <http://ts-dep.web.cern.ch/ts-dep/groups/ce/ce-LHC.htm>, June, 2012.
- [44] P. A. Aarnio, H. Abie, P. Abreu, and W. Adam, *The DELPHI detector at LEP*, Nucl. Instrum. Methods Phys. Res. , A **303** (1990) 233.

- [45] LHC Webpage, <http://lhc.web.cern.ch/lhc/>, June, 2012.
- [46] C. C. Group, *LHC The Guide*, CERN-Brochure-2008-001-Eng (January 2008).
- [47] The LHCb Collaboration, *LHCb : Technical Proposal*. Tech. Proposal. CERN, Geneva, 1998.
- [48] The LHCb Collaboration, *Review of the LHCb Reoptimized detector design and performance TDR*, Tech. Rep. CERN-LHCC-2003-064. LHCC-G-063, CERN, Geneva, Nov, 2003.
- [49] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008).
- [50] LHCb Collaboration, R. Aaij *et al.*, *Absolute luminosity measurements with the LHCb detector at the LHC*, JINST **7** (2012) P01010, [arXiv:1110.2866](https://arxiv.org/abs/1110.2866).
- [51] LHCb Detector Webpage, <http://lhcb-geom.web.cern.ch/lhcb-geom/>, June, 2012.
- [52] LHCb Webpage, <http://lhcb.web.cern.ch/>, June, 2012.
- [53] Amato, S *et al.*, *LHCb magnet: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [54] J. Andre *et al.*, *Status of the LHCb magnet system*, Applied Superconductivity, IEEE Transactions on **12** (2002) 366 .
- [55] J. Andre, W. Flegel, P. Giudici, O. Jamet, and M. Losasso, *Status of the lhcb dipole magnet*, Applied Superconductivity, IEEE Transactions on **14** (2004) 509 .
- [56] M. Losasso *et al.*, *Tests and field map of lhcb dipole magnet*, Applied Superconductivity, IEEE Transactions on **16** (2006), no. 2 1700.
- [57] J. Van Tilburg and M. Merk, *Track simulation and reconstruction in LHCb*. PhD thesis, Vrije Univ. Amsterdam, Amsterdam, 2005, Presented on 01 Sep 2005, CERN-THESIS-2005-040.
- [58] Adamus, M *et al.*, *LHCb VELO (VERtex LOCator): Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.
- [59] J. Gassner, F. Lehner, and S. Steiner, *The mechanical design of the lhcb silicon trigger tracker*, Tech. Rep. LHCb-2004-110. CERN-LHCb-2004-110, CERN, Geneva, Aug, 2010.
- [60] J. Gassner, F. Lehner, and S. Steiner, *The production, assembly and testing of the lhcb silicon trigger tracker*, Tech. Rep. LHCb-2004-109. CERN-LHCb-2004-109, CERN, Geneva, Jul, 2005.

- [61] Silicon Tracker Webpage, <http://lhcb.physik.unz.ch/SiliconTracker>, June, 2012.
- [62] Adamus, M et al., *LHCb Inner Tracker: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2002.
- [63] Adamus, M et al., *LHCb Outer Tracker: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.
- [64] Adamus, M et al., *LHCb RICH: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [65] P. A. Cherenkov, *Visible emission of clean liquids by action of gamma radiation*, Doklady Akademii Nauk SSSR **2** (1934) 451+.
- [66] Adamus, M et al., *LHCb Calorimeters: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [67] Adamus, M et al., *LHCb Muon System: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.
- [68] Adamus, M et al., *LHCb Muon System: addendum to the Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.
- [69] Adamus, M et al., *LHCb Muon System: second addendum to the Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2005.
- [70] G. Bencivenni et al., *A triple-GEM detector with pad readout for the inner region of the first LHCb muon station*, Tech. Rep. LHCb-2001-051, CERN, Geneva, Jul, 2001.
- [71] G. Graziani, R. Santacesaria, and A. Satta, *Study of the LHCb Muon Detector performance using 2010 beam data*, Tech. Rep. LHCb-PUB-2011-027. CERN-LHCb-PUB-2011-027, CERN, Geneva, Dec, 2011.
- [72] Adamus, M et al., *LHCb Trigger System: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.
- [73] R. Aaij and J. Albrecht, *Muon triggers in the High Level Trigger of LHCb*, Tech. Rep. LHCb-PUB-2011-017. CERN-LHCb-PUB-2011-017, CERN, Geneva, Sep, 2011.
- [74] V. V. Gligorov, *A single track HLT1 trigger*, Tech. Rep. LHCb-PUB-2011-003. CERN-LHCb-PUB-2011-003. LHCb-INT-2010-053, CERN, Geneva, Jan, 2011.
- [75] M. Williams et al., *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

- [76] F. Soomro and V. Belyaev, *HLT2 exclusive selections for  $B_s \rightarrow \phi\gamma$  and  $B_d \rightarrow K^*\gamma$* , Tech. Rep. LHCb-PUB-2010-007. CERN-LHCb-PUB-2010-007, CERN, Geneva, Feb, 2010.
- [77] P. M. S. Blackett and G. P. S. Occhialini, *Some photographs of the tracks of penetrating radiation*, Proceedings of the Royal Society of London, Series A **139** (1933) 699.
- [78] R. Mankel, *Pattern recognition and event reconstruction in particle physics experiments*, Rept. Prog. Phys. **67** (2004) 553, [arXiv:physics/0402039](#).
- [79] J. D. Jackson, *Classical Electrodynamics Third Edition*. Wiley, third ed., Aug., 1998.
- [80] S. Tavernier, *Experimental techniques in nuclear and particle physics*, Springer (2010).
- [81] ALICE Collaboration, K. Aamodt *et al.*, *Rapidity and transverse momentum dependence of inclusive  $J/\psi$  production in pp collisions at  $\sqrt{s} = 7$  TeV*, Phys. Lett. **B704** (2011) 442, [arXiv:1105.0380](#).
- [82] D. Groom, *Atomic and Nuclear Properties of Materials*, [pdg. lbl.gov/2011/AtomicNuclearProperties](#) (February 2011).
- [83] H. A. Bethe, *Molière's theory of multiple scattering*, Phys. Rev. **89** (1953) 1256.
- [84] G. R. Lynch and O. I. Dahl, *Approximations to multiple Coulomb scattering*, Nucl. Instrum. Meth. **B58** (1991) 6, Revised version.
- [85] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transactions of the ASME—Journal of Basic Engineering **82** (1960), no. Series D 35.
- [86] P. Billoir, *Track Fitting With Multiple Scattering: A New Method*, Nucl. Instrum. Meth. **A225** (1984) 352.
- [87] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.
- [88] E. Bos and E. Rodrigues, *The LHCb track extrapolator tools*, LHCb public note (2007).
- [89] M. Benayoun and O. Callot, *The forward tracking, an optical model method*, LHCb public note (2002).
- [90] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking*, LHCb public note (2007).
- [91] M. Needham and J. Van Tilburg, *Performance of the track matching*, LHCb public note (2007).



- [92] M. Needham, *Performance of the track matching*, LHCb public note (2007).
- [93] O. Callot and M. Schiller, *PatSeeding: A standalone track reconstruction algorithm*, LHCb public note (2008).
- [94] M. T. Schiller, *Track reconstruction and prompt  $K_S^0$  production at the LHCb experiment*. PhD thesis, Heidelberg University, 2011, Presented 12 Jul 2011, CERN-THESIS-2011-078.
- [95] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, LHCb public note (2011).
- [96] D. Hutchcroft, *VELO pattern recognition*, LHCb public note (2007).
- [97] O. Callot, *Downstream pattern recognition*, LHCb public note (2007).
- [98] O. Callot, M. Kucharczyk, and M. Witek, *VELO-TT track reconstruction*, LHCb public note (2007).
- [99] E. Rodrigues, *Dealing with clones in the tracking*, LHCb public note (2006).
- [100] A. Jaeger, P. Seyfert, M. De Cian, J. van Tilburg, and S. Hansmann-Menzemer, *Measurement of the track finding efficiency*, Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.
- [101] G. Manca, L. Mou, and B. Saitta, *Studies of Efficiency of the LHCb Muon Detector Using Cosmic Rays*, Tech. Rep. LHCb-PUB-2009-017. CERN-LHCb-PUB-2009-017. CERN-LHCb-INT-2009-026, CERN, Geneva, Oct, 2009.
- [102] G. Passaleva, *A recurrent neural network for track reconstruction in the lhc muon system*, Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE (2008) 867 .
- [103] G. Passaleva, *private communication*, .
- [104] P. Seyfert, *Measurement of Track Reconstruction Efficiencies at the LHCb Experiment*, Diploma thesis, Heidelberg University (2010).
- [105] O. Heaviside, XXXIX. *On the electromagnetic effects due to the motion of electrification through a dielectric*, Philosophical Magazine Series 5 **27** (1889), no. 167 324, arXiv:<http://www.tandfonline.com/doi/pdf/10.1080/14786448908628362>.
- [106] J. Gaiser, *Charmonium Spectroscopy from Radiative Decays of the  $J/\psi$  and  $\psi'$* . PhD thesis, SLAC, 1982.

- [107] J. Albrecht, *Fast Track Reconstruction for the High Level Trigger of the LHCb Experiment*. PhD thesis, Heidelberg, University of Heidelberg, Heidelberg, 2009, Presented on 08 Jul 2009, CERN-THESIS-2009-120.
- [108] G. Lanfranchi *et al.*, *The Muon Identification Procedure of the LHCb Experiment for the First Data*, Tech. Rep. LHCb-PUB-2009-013. CERN-LHCb-PUB-2009-013, CERN, Geneva, Aug, 2009.
- [109] CMS Collaboration, *Measurement of tracking efficiency*, CMS-PAS-TRK-10-002 (2010).
- [110] LHCb collaboration, R. Aaij *et al.*, *Inclusive  $W$  and  $Z$  production in the forward region at  $\sqrt{s} = 7$  TeV*, [arXiv:1204.1620](#).
- [111] M. Paterno, *Calculating efficiencies and their uncertainties*, , FERMILAB-TM-2286-CD.
- [112] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57** (1998) 3873.
- [113] A. Satta, *Muon identification in the LHCb high level trigger*, LHCb public note (2005).
- [114] BABAR Collaboration, B. Aubert *et al.*, *Evidence for the rare decay  $B \rightarrow K^* \ell^+ \ell^-$  and measurement of the  $B \rightarrow K \ell^+ \ell^-$  branching fraction*, Phys. Rev. Lett. **91** (2003) 221802, [arXiv:hep-ex/0308042](#).
- [115] BABAR Collaboration, B. Aubert *et al.*, *Angular Distributions in the Decays  $B \rightarrow K^* l^+ l^-$* , Phys. Rev. **D79** (2009) 031102, [arXiv:0804.4412](#).
- [116] Belle Collaboration, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for  $B \rightarrow K^{(*)} l^+ l^-$* , Phys. Rev. Lett. **103** (2009) 171801.
- [117] CDF Collaboration, T. Aaltonen *et al.*, *Measurements of the Angular Distributions in the Decays  $B \rightarrow K^{(*)} \mu^+ \mu^-$  at CDF*, Phys. Rev. Lett. **108** (2012) 081807, [arXiv:1108.0695](#).
- [118] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of  $B \rightarrow K^* \mu^+ \mu^-$  Decays in the Standard Model and Beyond*, JHEP **0901** (2009) 019, [arXiv:0811.1214](#).
- [119] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The Decay  $B \rightarrow K l^+ l^-$  at Low Hadronic Recoil and Model-Independent  $\Delta B = 1$  Constraints*, JHEP **1201** (2012) 107, [arXiv:1111.2558](#).
- [120] C. Bobeth, G. Hiller, and D. van Dyk, *More Benefits of Semileptonic Rare B Decays at Low Recoil: CP Violation*, JHEP **1107** (2011) 067, [arXiv:1105.0376](#).
- [121] M. Beneke, T. Feldmann, and D. Seidel, *Exclusive radiative and electroweak  $b \rightarrow d$  and  $b \rightarrow s$  penguin decays at NLO*, Eur. Phys. J. **C41** (2005) 173, [arXiv:hep-ph/0412400](#).

- [122] A. Ali, G. Kramer, and G.-h. Zhu,  $B \rightarrow K^+ l^+ l^-$  decay in soft-collinear effective theory, Eur. Phys. J. **C47** (2006) 625, [arXiv:hep-ph/0601034](#).
- [123] J. Matias, F. Mescia, M. Ramon, and J. Virto, *Complete Anatomy of  $\bar{B}_d \rightarrow \bar{K}^{*0}(\rightarrow K\pi)l^+l^-$  and its angular distribution*, JHEP **1204** (2012) 104, [arXiv:1202.4266](#).
- [124] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of  $B \rightarrow K^* l l$  at large recoil*, [arXiv:1207.2753](#).
- [125] LHCb Collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , Phys. Rev. Lett. **108** (2012) 181806, [arXiv:1112.3515](#).
- [126] LHCb Collaboration, LHCb collaboration, *Differential branching fraction and angular analysis of the  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  decay*, [arXiv:1304.6325](#).
- [127] M. Tresch, *Measurement of asymmetries in the decay  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , Master thesis, University of Zürich (2012).
- [128] D. Van Dyk, *The Decays  $B \rightarrow K^{(*)} \ell^+ \ell^-$  at Low Recoil and their Constraints on New Physics*. PhD thesis, TU Dortmund, 2012.
- [129] Particle Data Group, K. Nakamura *et al.*, *Review of particle physics*, J. Phys. G **G37** (2010) 075021.
- [130] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, Tech. Rep. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, CERN, Geneva, Jan, 2008.
- [131] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](#).
- [132] C. Salzmann, *LHCb Spectrometer Alignment and Verification of its Performance using the Decay  $B_d^0 \rightarrow K^{*0} J/\psi$* . PhD thesis, University of Zurich, 2011, Presented 08 Dec 2011, CERN-THESIS-2011-201.
- [133] J. Dickens, V. Gibson, C. Lazzeroni, and M. Patel, *Selection of the Decay  $B_d \rightarrow K^{*0} \mu^+ \mu^-$  at LHCb*, Tech. Rep. LHCb-2007-038. CERN-LHCb-2007-038, CERN, Geneva, Apr, 2007.
- [134] F. Jansen, N. Serra, G. Y. Smit, and N. Tuning, *Determination of the forward-backward asymmetry in the decay  $B^0 \rightarrow K^* \mu \mu$  with an unbinned counting analysis*, Tech. Rep. LHCb-2009-003. CERN-LHCb-2009-003, CERN, Geneva, May, 2009.
- [135] B. Efron, *Bootstrap methods: Another look at the jackknife*, The Annals of Statistics **7** (1979), no. 1 1.

- [136] B. Aslan and G. Zech, *Statistical energy as a tool for binning-free, multivariate goodness-of-fit tests, two-sample comparison and unfolding*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **537** (2005), no. 3 626 .
- [137] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST **5** (2010) P09004, [arXiv:1006.3019](#).
- [138] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of  $B \rightarrow K^* \ell^+ \ell^-$  observables in the full kinematic range*, JHEP **1305** (2013) 137, [arXiv:1303.5794](#).
- [139] J. Matias, *On the S-wave pollution of  $B \rightarrow K^* \ell^+ \ell^-$  observables*, [arXiv:1209.1525](#).
- [140] D. Becirevic and A. Tayduganov, *Impact of  $B \rightarrow K_0^* \ell^+ \ell^-$  on the New Physics search in  $B \rightarrow K^* \ell^+ \ell^-$  decay*, [arXiv:1207.4004](#).

# Acknowledgements

All good things must come to an end and this thesis is no exception. It is thanks to many people that the time of my PhD was very enjoyable.

I am very grateful to Ueli Straumann for giving me the opportunity to do a PhD in his group. I would like to thank him for giving me a lot of freedom in my PhD research and letting me pursue some independent little projects throughout my time at the university. And also for sharing some of his vast knowledge about everything related to railways.

Many thanks go to Olaf Steinkamp for good advice throughout my whole PhD in many aspects — and certainly knowing more about the perfect presentation than anyone could have thought.

My thanks then go to the «islanders», Jonny Anderson and Mark Tobin. I am thankful to Jonny for supervising me in the first part of my PhD and for his happy approach to physics. His «quick cross-checks» will always be remembered. I am grateful to Mark for carefully proofreading my thesis and polishing its English, teaching me weird bash-commands and introducing the office to the wonderful world of British humour. Scorchio!

Many thanks also go to Jeroen van Tilburg, for many discussions about tracking and proof-reading the tracking parts of my thesis, and to Katharina Müller for sharing her large experience in dealing with all aspects of particle physics.

I am indebted to «Our Man in Geneva», Nicola Serra, for supervising the analysis part of this thesis and teaching me (almost) everything I ever wanted to know about flavour physics. Our common excursions into the realm of form factor uncertainty free observables with an ever changing naming scheme, discussed in numerous Skype conversations, certainly form a cornerstone of the work presented in this thesis. And I also thank him for his many new ideas and his loyalty.

Further thanks go to Christophe Salzmann for our daily (weekly) common rants (discussions) about politics, land-use planning and the best usage of frequent flyer miles. The introduction into horrendously written Brunel code and the key to not getting mad at it was much appreciated. Thanks also go to Angela Büchler for her invaluable contributions to «Mad Fridays»; and to Albert Bursche for sharing his python knowledge and his musical preferences, reminding me of the time when «Happy Hardcore» roamed the surface of this planet.

I am thankful to the «crew of the other Straumann office» for not minding (too much) my periodic outbursts of anger or excitement about anything and nothing. My thanks go to Christian, Marco and Nicola for many interesting and fruitful discussions, may it be about physics or other important things in life. And of course for the discovery of an uncountable

number of things of paramount importance on the world wide web. Many thanks also go to Ben and Arno for having a good time and an unforgettable trip to the US, or at least the plans for it. Thanks also to Roman for good advice in computing issues and the exploration of vortices of pure procrastination.

Further thanks go to Roland Bernet for figuring out answers to computing questions while I was still formulating them; Achim, Daniel and Stefan for helping with everything hardware / electronics related (and my theremin!); Carmelina for an occasional cookie break; and Monika and Ruth for always making sure the money arrived at my bank account, still leaving me to wonder where it actually came from in the first place.

My thanks then also go to all the people from various institutes throughout the world who helped me and my projects become reality.

I am very grateful to all my friends and to my parents who supported me during the last four years.

# Curriculum Vitae

## Personal information:

Last name	De Cian
First name	Michel
Day of Birth	31 <sup>st</sup> October 1982
Hometown	Brügg, BE

## Education:

2008 - 2013	PhD in particle physics, University of Zurich Title of PhD thesis: <i>Track Reconstruction Efficiency and Analysis of <math>B^0 \rightarrow K^{*0} \mu^+ \mu^-</math> at the LHCb Experiment</i> Supervisor: U. Straumann
2008	Internship as a science writer at the «SonntagsZeitung»
2003 - 2008	Diploma studies of physics, ETH Zurich Title of diploma thesis: <i>The search for <math>B_s^0 \rightarrow \mu \tau</math> at CMS</i> Supervisor: U. Langenegger
1999 - 2003	Matura at Alte Kantonsschule Aarau Schwerpunktfach: <i>Physik &amp; Anwendungen der Mathematik</i> Ergänzungsfach: <i>Philosophie</i>

This thesis was typeset in L<sup>A</sup>T<sub>E</sub>X using Garamond and Maven Pro.